AD 414717

FTD-TT-62-1367

TRANSLATION

ARITHMETIC SIMULATION OF RANDOM PROCESSES

By

A. G. Postnikov

FOREIGN TECHNOLOGY DIVISION

AIR FORCE SYSTEMS COMMAND

WRIGHT-PATTERSON AIR FORCE BASE

OHIO

FTD-TT-62-1367/1+2

UNEDITED ROUGH DRAFT TRANSLATION

ARITHMETIC SIMULATION OF RANDOM PROCESSES

BY: A. G. Postnikov

English Pages: 98

(Ref. in) S/44-62-0-1-3-61

FTD-TT-62-1367/1+2 Date 4 June 1963

S

Akademiya Nauk

Soyuza Sovetskikh Sotsialisticheskikh Respublik

ARIFMETICHESKOYE MODELIROVANYYE SLUCHAYNYKH PROTSESSOV

Trudy Matematicheskogo Instituta
Imeni V. A. Steklov

Vol. 57

Izdatel'stvo Akademii Nauk SSSR
Moskva - 1960

Pages: 1-84

FTD-TT- [illegible]

TABLE OF CONTENTS

Introduction 1

§ 1. Normal Sequence of Symbols 10

§ 2. Champernowne's Example of a Normal Sequence of Terms . 10

§ 3. Application to Uniform Distribution of Fractional Parts of an Exponential Function 15

§ 4. Criteria for Normality of Base Sequence of Terms . . . 16

§ 5. Application of the Criterion of Normality for Sequences of Terms 21

§ 6. A Second Definition of the Normal Sequence of Symbols. 22

§ 7. A System of Mutually Normal Sequences of Terms 25

§ 8. The Problem of Constructing a Sequence Mutually Normal to a Given Sequence 27

§ 9. Bernoulli-Normal Sequences of Terms 30

§ 10. Construction of a Bernoulli-Normal Sequence of Symbols 35

§ 11. Relationship of the Concepts of Bernoulli-Normal Sequence of Symbols and Admissible Number 38

§ 12. Completely Uniformly Distributed Sequences 41

§ 13. Construction of a Completely Uniformly Distributed Sequence 43

§ 14. N.M. Korobov's Theorem 46

§ 15. The Converse of N.M. Korobov's Theorem 46

§ 16. A Sequence Completely Distributed Over the Function F(x) 48

§ 17. Construction of a Sequence Completely Distributed With Respect to a Function F(x) 50

§ 18. Measure in a Space of Infinite Sequences of Symbols . 51

§ 19. A Dynamic System in a Space of Sequences of Terms . . 57

§ 20. Birkhoff-Khinchin Theorem 59

§ 21. Proof of the Normality Criterion for Sequences of Symbols on the Basis of the Birkhoff Theorem 67

§ 22. Dynamic System Corresponding to the Simplest Markov Chain 70

§ 23. Markov-Normal Sequences . . . . 7[illegible]

§ 24. Construction of Markov-Normal Sequence of Symbols . . . 74
§ 25. The Dynamic System in the Theory of Continued Fractions . . . 77
§ 26. Normal Continued Fraction . . . 81
§ 27. Construction of a Normal Continued Fraction . . . 82

Notes . . . 88

References . . . 91

The author dedicates this work
in respectful memory of
Aleksandr Yakovlevich Khinchin

INTRODUCTION

Problems on Diophantine approximations with exponential functions form one of the sources for the material presented in this monograph.

Let $g \geq 2$ be a natural number. Let α be a real number, $0 \leq \alpha \leq 1$. Let us consider the sequence of fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots$. We let δ be an interval on the segment [01], and $|\delta|$ its length. We let $N_p(\delta)$ be the number of fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots, P$ lying in the interval δ. We shall say that a sequence of fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots$, is uniformly distributed on [01] if the relationship

$$\lim_{P\to\infty}\frac{N_P(\delta)}{P}=|\delta|.$$

is valid for any interval δ.

Let us consider the infinite sequence consisting of the terms 0, 1, ..., $g - 1$,

$$a_1, a_2, a_3, \ldots . \qquad (1)$$

We select a natural number s and write a sequence of s-tuples

$$(a_1 a_2 \ldots a_s)(a_2 a_3 \ldots a_{s+1})(a_3 \ldots a_{s+2}) \ldots \qquad (2)$$

Let Δ be any fixed s-tuple consisting of the terms 0, 1, ..., $g - 1$. We let $N_p(\Delta)$ be the number of times the s-tuple Δ is encountered prior to the Pth term of the Sequence (2).

We shall call the sequence (1) a normal sequence if

$$\lim_{P\to\infty}\frac{N_P(\Delta)}{P}=\frac{1}{g^s}.$$

for any natural number s and any s-tuple.

We decompose α into an infinite fraction written to the base g

$$\alpha=\frac{a_1}{g}+\frac{a_2}{g^2}+\frac{a_3}{g^3}+\cdots \qquad (3)$$

A well-known theorem ([1], page 233) states: a uniform distribution of the fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots$ on the segment [01] is equivalent to the normality of the sequence

$$a_1, a_2, a_3, \ldots \quad (4)$$

Problems of Diophantine approximations with exponential functions have been the subject of careful study. One of the first results to appear was the theorem stating that with respect to the Lebesque measure for almost all numbers α, $0 \leq \alpha \leq 1$, the sequence of fractions $\{\alpha g^x\}$ $x = 1, 2\ldots$ is uniformly distributed on the segment [01] (see [2]). Thus the existence of normal sequences of digits was established.

Borel took a similar approach. Borel ([3], page 197) called a real number α, $0 \leq \alpha \leq 1$ weakly normal with respect to a base $\underline{g}$ (I have deliberately translated "simplement normal" as weakly normal) if the sequence

$$a_1, a_2, a_3, \ldots,$$

obtained by decomposing α into an infinite fraction written to base $\underline{g}$ $\alpha = (a_1/g) + (a_2/g^2) + \ldots$, has the property that each of the terms appearing in it occurs with an asymptotic frequence equal to $1/g$. Moreover, Borel calls a number α absolutely normal if it is weakly normal with respect to every natural base $\underline{g}$ larger than unity.

On the basis of measure theory, Borel established the existence of absolutely normal numbers.

In §6 of the present monograph, it is proved that an absolutely normal number α possesses the property that no matter what natural base $\underline{g}$ is used in the decomposition of α,

$$a_1 = \frac{a_1}{g} + \frac{a_2}{g^2} + \cdots,$$

the sequence a_1, a_2, ... will be a normal sequence and, consequently, for any natural $g \geq 2$ the fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots$, will be normally distributed.

In our monograph, the emphasis has been shifted from problems involving the distribution of fractional portions of an exponential function to the normal sequences.

Problems associated with the use of the words "table of random numbers" or "table of pseudorandom numbers" also furnish some of the material discussed in this monograph. The discussion contains the stipulation that we are speaking of a table that has no bound in one direction, i.e., we are concerned with an infinite sequence of numbers.

Certain authors do not define these words when they use them. These authors include persons concerned with the practical utilization of such tables. Kendall and Smith ([4], page 167) in an article concerned with tests for checking numerical sequences for "randomness" write that "... for the purposes of this article, the logical aspect has been deemphasized... ." In his report, Steinhaus [5] notes that the words "random sequence" are in daily use by statisticians, although they do not define these words. In particular, such lack of precision may be found in several studies dealing with the Monte Carlo method ([6] for example).

We find a desire to employ these words precisely in Venn [7] who is of the opinion that "randomness" should be defined in terms of frequency.

Mises follows Venn ([8], page 28). Mises has introduced the word collective. A collective is defined by two conditions ([8], page 31).

1. The relative frequencies of any terms must have definite limiting values.

2. The limiting frequency value of a term "should remain invariant if any portion of the sequence is arbitrarily selected and

just this portion is then examined" (requirement of collective irregularity).

The formulation of Mises' second condition is not clear. To see how to give Mises' hints a precise meaning, see [9], page 218, [10], and [60].

Some authors departed from Mises' program to isolate subsequences: they remained on the path described by Venn and grouped terms in the sequence, studying their distribution. In particular, this was the approach used by Copeland [11] who introduced the concept of the admissible number (see §11), by Reichenbach, who introduced a concept identical with the concept of the admissible number [61], by A.G. Postnikov and I.I. Pyatetskiy [12], who introduced the concept of a Bernoulli-normal sequence.

Let there be given two positive numbers p and q such that p + q = = 1. Consider the infinite sequence composed of the symbols 0 and 1,

$$\varepsilon_1 \varepsilon_2 \varepsilon_3 \ldots \tag{5}$$

Let s be any natural number. We write Sequence (5) as a "caterpillar."

$$(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_s)(\varepsilon_2 \varepsilon_3 \ldots \varepsilon_{s+1}) \ldots (\varepsilon_P \varepsilon_{P+1} \ldots \varepsilon_{P+s-1}) \ldots \tag{6}$$

Let $\Delta = (\delta_1 \ldots \delta_s)$ be any s-tuple consisting of the symbols 0 and 1. We let $N_P(\Delta)$ be the number of appearances of the tuple Δ prior to the Pth term of Sequence (6). We call Sequence (5) a Bernoulli-normal sequence of symbols if for any natural s and any s-tuple,

$$\lim_{P \to \infty} \frac{N_P(\Delta)}{P} = p^l q^{s-l},$$

where l is the number of ones among the symbols of $\Delta = (\delta_1 \ldots \delta_s)$.

The strong law of large numbers for stationary random sequences (see [13], page 417) permits us to establish, in particular, this theorem.

Theorem. Let an unlimited number of independent trials be carried

out; for each trial the event designated by the symbol 1 has probability p, while the event designated by the symbol 0 has probability q. With probability unity, the sequence of outcomes will be Bernoulli-normal.

This theorem, in particular, establishes the existence of Bernoulli-normal sequences of symbols.

In §11 it is proved that the concepts of admissible number and Bernoulli-normal sequence of symbols are equivalent. The introduction of the concept of Bernoulli-normal sequences while making more precise the use of the term "table of pseudorandom numbers" is justified by the fact that such sequences of numbers (or more general sequences) are clearly sufficient for the construction of a numerical method of analysis similar to the Monte Carlo method, but yielding a reliable error estimate (see [14]).

The reader will find additional material on the problem of tables of random numbers in the Supplement(1).

The concept of a sequence fully distributed with respect to a function F(x) represents a generalization of the concept of a normal sequence of symbols and the concept of a Bernoulli-normal sequence of symbols.

Let there be given a distribution function F(x). Consider the infinite sequence of real numbers

$$a_1, a_2, a_3, \ldots . \tag{7}$$

We choose any natural number s and any system $\Delta = (\Delta_1, \ldots, \Delta_s)$ of intervals $\Delta_1 = (a_1 b_1), \ldots, \Delta_s = (a_s b_s)$, whose end points are points of continuity for the function F(x). We form the line

$$(a_1 a_2 \ldots a_s)(a_2 a_3 \ldots a_{s+1}) \ldots (a_P a_{P+1} \ldots a_{P+s-1}) \ldots \tag{8}$$

and let $N_P(\Delta)$ be the number of tuples prior to the Pth term of Sequence (8) in which the first component belongs to Δ_1, the second to Δ_2, ..

and the sth to Δ_s. We say that Sequence (7) is completely distributed with respect to the function $F(x)$ if for any s and any system of intervals Δ the asymptotic relationship

$$\lim_{P\to\infty}\frac{N_P(\Delta)}{P}=(F(b_1)-F(a_1))\dots(F(b_s)-F(a_s)). \tag{9}$$

is valid.

A special case of this concept – the concept of a completely uniformly distributed sequence in which $F(x) = x$, $0 \leq x \leq 1$ was introduced by N.M. Korobov [1] (in another form).

The concept of a normal sequence of symbols is obtained from this general concept when the sequence consists of the numbers 0, 1, ..., $g - 1$, and the distribution $F(x)$ equals

$$F(x)=\begin{cases}0 & x<0\\ \frac{1}{g} & 0<x<\frac{1}{g}\\ \frac{2}{g} & \frac{1}{g}<x<\frac{2}{g}\\ \dots\dots\dots & \\ 1 & 1<x.\end{cases}$$

The concept of a Bernoulli-normal sequence of symbols is obtained from this concept when the sequence consists of the numbers 0 and 1, and the distribution function $F(x)$ equals

$$F(x)=\begin{cases}0 & x<0\\ q & 0<x<1\\ 1 & 1<x.\end{cases}$$

For any distribution function $F(x)$ there exists a sequence that is completely distributed over the function $F(x)$. This follows from the strong law of large numbers for stationary random sequences ([13], page 417), which in this special case yields the following generalization of Glivenko's theorem ([15], page 328).

Theorem. Let the random variable ξ have a distribution function $F(x)$. Let us take an infinite sample of this random variable:

$$a_1, a_2, \dots, a_P, \dots$$

This sequence is completely distributed with respect to the func-

tion F(x) with probability equal to unity.

I next introduce a concept broader than the concept of a sequence completely distributed over the function F(x).

Let there be given a random sequence stationary in the narrow sense, i.e., a sequence of random variables

$$\xi_1, \xi_2, \xi_3, \ldots \quad (10)$$

such that for any set of natural numbers $n_1 < n_2 < n_s$, any set of intervals Δ_1, Δ_2, ..., Δ_s on the real line, and any natural n

$$P(\xi_{n_1} \in \Delta_1, \ldots, \xi_{n_s} \in \Delta_s) = P(\xi_{n_1+n} \in \Delta_1, \ldots, \xi_{n_s+n} \in \Delta_s)$$

(P is the probability).

We assume that Sequence (10) is metrically transitive (see [13], page 410).

Let there be given the infinite sequence of real numbers

$$\alpha_1, \alpha_2, \alpha_3, \ldots . \quad (11)$$

Let us take any natural number s and any system $\Delta = (\Delta_1, \ldots, \Delta_s)$ of s intervals on the real line. We write

$$(\alpha_1\alpha_2 \ldots \alpha_s)(\alpha_2\alpha_3 \ldots \alpha_{s+1}) \ldots \quad (12)$$

and let $N_P(\Delta)$ be the number of times the s-tuple in which

$$\alpha_i \in \Delta_1, \alpha_{i+1} \in \Delta_2, \ldots, \alpha_{i+s-1} \in \Delta_s.$$

is encountered prior to the Pth term of Sequence (12).

The infinite Sequence (11) is called a normal realization of a stationary random sequence (the term was proposed by A.N. Kolmogorov) if for any natural s and any system of intervals $\Delta = (\Delta_1, \ldots, \Delta_s)$ the equality

$$\lim_{P \to \infty} \frac{N_P(\Delta)}{P} = P(\xi_1 \in \Delta_1, \ldots, \xi_s \in \Delta_s). \quad (13)$$

is valid.

On the basis of the strong law of large numbers for stationary random sequences (see [13], page 417) a realization of a metrically transitive stationary random sequence will be normal with probabi[lity]

equal to unity.

The problem of constructing normal realizations of stationary random processes by arithmetic means is of interest. A sequence constructed by arithmetic means is called a sequence given with the aid of a primitive recursion function.

Sierpinski [16] dealt with an effective definition of absolutely normal numbers. Lebesgue's study [17] belongs to this group of problems.

Champernowne [18] constructed a normal sequence of symbols (he constructed a Bernoulli-normal sequence with $p = q = 1/2$, and $g = 2$). Other methods for constructing normal sequences of symbols have been given by Copeland and Erdős [50] and Davenport and Erdős [51].

The argument used to prove that a sequence written by Champernowne's method is a normal sequence of symbols is quite complicated. A.G. Postnikov [19] has noted that the argument is simplified considerably if the following criteria established by I.I. Pyatetskiy [20] are used.

Theorem. Let there be a sequence consisting of the symbols

$$0, 1, \ldots, g-1,$$
$$a_1, a_2, a_3, \ldots \tag{14}$$

such that there exists a constant $C > 0$ such that for any natural number s and any s-tuple

$$\varlimsup_{p\to\infty} \frac{N_p(\Delta)}{p} < \frac{C}{g^s}.$$

Then Sequence (14) is a normal sequence of symbols.

A.G. Postnikov and I.I. Pyatetskiy [12] have extended Champernowne's method and have constructed a Bernoulli-normal sequence of symbols for arbitrary p. Here a theorem similar to the criterion mentioned above was used.

In like manner, A.G. Postnikov and I.I. Pyatetskiy [21] constructed a normal realization of a very simple stationary Markov chain.

Finally, A.G. Postnikov and I.I. Pyatetskiy [21] also constructed the normal realization of a process corresponding to a continued fraction.

N.M. Korobov [1] has suggested a method for constructing normal sequences of symbols* that is based on the utilization of normal periodic systems (the concept of normal periodic systems was evidently first introduced by Martin [22]. Yu.N. Shakhov [24] used a generalization of this concept suggested by N.M. Korobov [23] to solve the problem of constructing a normal realization of a very simple Markov chain (Yu.N. Shakhov imposes stronger limitations on the transition probabilities than are imposed in [21]). We shall not be concerned with the method of normal periodic systems in this monograph.

N.M. Korobov [1, 25] used different methods to construct completely uniformly distributed sequences. L.P. Starchenko [27] succeeded in constructing a completely uniformly distributed sequence.

With the aid of a completely uniformly distributed sequence, N.M. Korobov solved the problem of constructing a normal sequence of symbols [1], and gave a multidimensional generalization of this problem [25]. As we shall show in this monograph (§17) a completely uniformly distributed sequence may be used to construct a sequence that is completely distributed with respect to a function $F(x)$.**

The aim of the present study is to examine a portion of the material that has accumulated in this field from the viewpoint mentioned above.

I wish to thank all those who helped me in this work.

§1. Normal Sequence of Symbols

Consider an infinite sequence composed of the terms 0, 1, ..., g − 1,

$$\alpha = s_1 s_2 \ldots s_P \ldots \qquad (1)$$

We take any natural number s and construct a sequence of s-tuples

$$(s_1 s_2 \ldots s_s)(s_2 s_3 \ldots s_{s+1}) \ldots (s_P s_{P+1} \ldots s_{P+s-1}) \ldots . \qquad (2)$$

We call Sequence (2) a caterpillar (of rank s) of Sequence (1). Let $\Delta = (\delta_1 \ldots \delta_s)$ be any s-tuple (we shall also use the word combination) consisting of the symbols 0, 1, ..., g − 1. We let $N_P(\alpha, \Delta)$ or $N_P(\Delta)$ be the number of occurrences of the s-tuple Δ prior to the Pth term of Sequence (2).

Definition. We call Sequence (1) a normal sequence of symbols if for any natural number s and any s-tuple Δ the limiting relationship

$$\lim_{P \to \infty} \frac{N_P(\Delta)}{P} = \frac{1}{g^s}.$$

is satisfied.

§2. Champernowne's Example of a Normal Sequence of Terms

Following Champernowne [18] let us construct a normal sequence of symbols.

We let s_r be a sequence consisting of all r-digit numbers written in the g scale; here we also consider a combination of symbols beginning with a zero to be an r-digit number. We take the numbers in their natural order. For example where g =

$$s_1 = 0'1$$
$$s_2 = 00'01'10'1.$$
$$s_3 = 000'001'010'011'100'101'110'111$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$

In writing the s_r, we will place apostrophes between the r-digit numbers, as shown. Let us prove that the sequence

$$s_1 s_2 s_3 \ldots s_r \ldots$$

is normal (since s_r is a group of terms, and not a single term, the notation is symbolic). We must show that each ρ-digit combination of symbols (ρ is fixed) Δ_ρ in the caterpillar of Sequence (1) is encountered with an asymptotic frequency $1/g^\rho$.

We let:

S_r be the sequence $s_1 s_2 \ldots s_r$;

x_r be the number of base-g terms in s_r;

X_r be the number of base-g terms in S_r;

g_r be the number of appearances of Δ_ρ in s_r;

G_r be the number of appearances of Δ_ρ in S_r;

$g_r(x)$ be the number of appearances of Δ_ρ among the first <u>x</u> terms in s_r;

$G(x)$ be the number of appearances of Δ_ρ among the first <u>x</u> terms of Sequence (1).

Let $\Delta_\rho = (\delta_1 \delta_2 \ldots \delta_\rho)$.

We must show that

$$G(x) = \frac{x}{g^\rho} + o(x).$$

If Δ_ρ enters into a sequence s_r so that apostrophes do not separate its terms, we shall say that Δ_ρ enters undivided; if an apostrophe separates the terms of Δ_ρ we shall then say that Δ_ρ enters divided. For example, when $g = 2$, $\Delta_3 = (101)$ enters s_3 undivided in 100' 101' 110' and divided in 110' 111.

If $r < \rho$, Δ_ρ cannot be contained undivided in s_r. If $r \geq \rho$ then Δ_ρ is contained undivided in s_r exactly $(r - \rho + 1)\, g^{r-\rho}$ times. Actually, there are $r - \rho + 1$ ways in which Δ_ρ may enter undivided into s_r: The first term of Δ_ρ may be the first term of an r-digit number, the second term, etc. If we select a position for Δ_ρ, we may take all the remaining $r - \rho$ terms arbitrarily. Thus Δ_ρ is encountered in s_r

a combination undivided by apostrophes exactly $(r-\rho+1)\,g^{r-\rho}$ times. What is more, s_r contains g^r apostrophes. When an apostrophe is given, it cannot separate more than ρ different Δ_ρ. As a result, no Δ_ρ divided more than ρg^r times can be found in s_r. Thus

$$g_r = (r-\rho+1)\,g^{r-\rho} + O(g^r),$$

where $r \to \infty$.

But

$$x_r = r g^r.$$

Consequently

$$g_r = \frac{x_r}{g^\rho} + o(x_r).$$

Moreover

$$G_r = \sum_{s=1}^{r} g_s + O(r),\quad X_r = \sum_{s=1}^{r} x_s.$$

Thus

$$G_r = \frac{X_r}{g^\rho} + o(X_r).$$

Let us evaluate $g_r(x)$. We assume that $\underline{x}$ is found in the number $'p_{r-1}p_{r-2}\dots p_1p_0$ of sequence s_r. It is clear that

$$x = r\sum_{i=0}^{r-1} p_i g^i + \theta r,\quad 0<\theta\leq 1.$$

We recall that Δ_ρ in any r-digit number s_r may occupy $r-\rho+1$ different positions. We let $g_{rk}(x)$ be the number of appearances of Δ_ρ in undivided form among the first $\underline{x}$ terms of s_r in a position such that the first term of Δ_ρ coincides with the <u>kth</u> digit of the number. If $k > r-\rho+1$ such a position does not exist and $g_{rk}(x) = 0$. We shall prove that where $k \leq r-\rho+1$

$$g_{rk}(x) = g^{r-\rho+1-k}\left(\sum_{i=r-k+1}^{r-1} p_i g^{i-(r-k+1)} + \theta'\right),\quad 0<\theta'<1$$

(for the case $k = 1$, the sum vanishes).

The structure of an r-digit number s_r in which the first term of

Δ_ρ coincides with the kth digit of this number may be represented as follows:

$$\underbrace{\qquad\qquad}_{k-1\,\text{terms}}\delta_1\delta_2\ldots\delta_\rho\underbrace{\qquad\qquad\qquad}_{r-\rho+1-k\,\text{terms}}.$$

We now count these numbers up to (and including) the number

$$'\underbrace{p_{r-1}\ldots p_{r-(k-1)}}_{k-1\,\text{terms}}\ \underbrace{p_{r-k}\ldots p_{r-(k+\rho-1)}}_{\rho\,\text{terms}}\ \underbrace{\qquad\qquad p_0'}_{r-\rho+1-k\,\text{terms}}.$$

1) If $p_{r-k}g^{\rho-1} + p_{r-k-1}g^{\rho-2} + \ldots + p_{r-(k+\rho-1)} > \delta_1 g^{\rho-1} + \delta_2 g^{\rho-2} + + \ldots + \delta_\rho$ it is evident that the quantity sought will equal the amount of these numbers up to and including the number

$$'p_{r-1}\ldots p_{r-(k-1)}\delta_1\delta_2\ldots\delta_\rho\, g-1\ldots g-1'.$$

It is clear that the last $r-\rho+1-k$ terms may be taken arbitrarily. The first $k-1$ terms may be selected in

$$\sum_{i=r-(k-1)}^{r-1} p_i g^{i-(r-k+1)} + 1$$

ways. Thus in this case the quantity desired will equal

$$g^{r-\rho+1-k}\left(\sum_{i=r-(k-1)}^{r-1} p_i g^{i-(r-k+1)} + 1\right).$$

2) If $p_{r-k}g^{\rho-1} + p_{r-k-1}g^{\rho-2} + \ldots + p_{r-(k+\rho-1)} < \delta_1 g^{\rho-1} + \delta_2 g^{\rho-2} + + \ldots \delta_\rho$, the desired quantity will equal the amount of these numbers up to and including the number

$$'p_{r-1}\cdots p_{r-(k-2)}\,p_{r-(k-1)}-1\,\delta_1\delta_2\cdots\delta_\rho\, g-1\cdots g-1'.$$

Thus in this case the number sought will equal

$$g^{r-\rho+1-k}\sum_{i=r-(k-1)}^{r-1} p_i g^{i-(r-k+1)}.$$

3) If $p_{r-k}g^{\rho-1} + p_{r-k-1}g^{\rho-2} + \ldots + p_{r-(k+\rho-1)} = \delta_1 g^{\rho-1} + \ldots + \delta_\rho$ the quantity of these numbers will equal the number of numbers up to

$$'p_{r-1}\cdots p_{r-(k-2)}\,p_{r-(k-1)}-1\,\delta_1\delta_2\cdots\delta_\rho\, g-1\ g-1\cdots g-1$$

plus an amount not exceeding $g^{r-(k+\rho-1)}$ since not all combinations may enter into the last $r-(k+\rho-1)$ terms. Thus, in all cases

$$g_{rk}(x) = g^{r-\rho+1-k}\left(\sum_{t=r-(k-1)}^{r-1} p_t g^{t+k-r-1} + \theta'\right), \quad |\theta'| < 1.$$

Moreover

$$\sum_{k=1}^{r-\rho+1} g_{rk}(x) = \sum_{k=1}^{r-\rho+1} \sum_{t=r-(k-1)}^{r-1} p_t g^{t-\rho} + \theta' \sum_{k=1}^{r-\rho+1} g^{r-\rho-k+1} =$$
$$= \sum_{t=\rho}^{r-1} p_t g^{t-\rho}(t+1-\rho) + O(g^r).$$

In view of the fact that there are no more than $O(g^r)$ divided Δ_r, we obtain

$$g_r(x) = \frac{1}{g^\rho}\sum_{t=\rho}^{r-1}(t+1-\rho)p_t g^t + O(g^r).$$

But

$$x = r\sum_{t=0}^{r-1} p_t g^t + \theta r,$$

$$\frac{x}{g^\rho} - g_r(x) = \frac{\sum_{t=\rho}^{r-1}(r+\rho-t-1)p_t g^t + r\sum_{t=0}^{\rho-1} p_t g^t}{g^\rho} + O(g^r) =$$

$$= \frac{\sum_{t=\rho}^{r-1}(r+\rho-t-1)p_t g^t}{g^\rho} + O(g^r) = O\left(g^r \sum_{t=\rho}^{r-1} \frac{r+\rho-t-1}{g^{r-t}}\right) + O(g^r) = O(g^r)$$

[in view of the convergence of the series $\sum_{t=1}^{\infty}\frac{t}{g^t}$].

Thus

$$g_r(x) = \frac{x}{g^\rho} + o(x) \quad \text{for } r \to \infty.$$

Let the <u>xth</u> term of Sequence (1) be the <u>yth</u> term in s_r. Then

$$x = X_{r-1} + y.$$
$$G(x) = G_{r-1} + g_r(y) + O(1), \text{ for } x \to \infty.$$

Hence

$$G(x) = \frac{X_{r-1}}{g^\rho} + \frac{y}{g^\rho} + O(g^r),$$

$$G(x) = \frac{x}{g^\rho} + o(x),$$

which has to be proved.

§3. Application to Uniform Distribution of Fractional Parts of an Exponential Function

Let $g \geq 2$ be an integer, and α a real number.

Let us examine the sequence of fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots, P, \ldots$. Let Δ be some half-interval on the segment [01], mes δ the length of this half interval. Let

$$\alpha = \frac{\varepsilon_1}{g} + \frac{\varepsilon_2}{g^2} + \cdots \tag{1}$$

be a decomposition of α written to base g (we assume that α is irrational, and thus the base-g expansion is uniquely determined.

Theorem 1. A necessary and sufficient condition for the fraction $\{\alpha g^x\}$ to lie on a half interval of the form $\left[\frac{a}{g^s}, \frac{a+1}{g^s}\right)$, where s is any integer ≥ 1, and a is an integer, $0 \leq a \leq g^s - 1$, $a = \delta_1 g^{s-1} + \ldots + \delta_s$, $(0 \leq \delta_i \leq g - 1)$, is the presence of the term $(\delta_1 \ldots \delta_s)$ at the xth position in the caterpillar of rank s of the sequence $\varepsilon_1 \varepsilon_2 \varepsilon_3 \cdots$ (2)

Proof

$$\{\alpha g^x\} = \frac{\delta_1}{g} + \frac{\delta_2}{g^2} + \cdots + \frac{\delta_s}{g^s} + \theta\left(\frac{g-1}{g^{s+1}} + \frac{g-1}{g^{s+2}} + \cdots\right),$$

where $0 \leq \theta < 1$. Thus

$$0 < \{\alpha g^x\} - \frac{a}{g^s} < \frac{1}{g^s}.$$

We let $N_P(\delta)$ be the number of fractions among the $\{\alpha g^x\}$, $x = 1, 2, \ldots, P$, lying within δ. We say that the sequence of fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots$, is uniformly distributed on the segment [01] if no matter what half interval δ on [01] we choose the number $N_P(\delta)$ satisfies the asymptotic relationship

$$\lim_{P \to \infty} \frac{N_P(\delta)}{P} = \text{mes } \delta.$$

when $P \to \infty$.

Theorem 2. If the fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots, P, \ldots$, where α is defined by Equation (1), are uniformly distributed, then the ba-

g terms of the number α in (2) form a normal sequence of symbols. Conversely, if Sequence (2) is a normal sequence of symbols, then the fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots, P, \ldots$, where

$$\alpha = \frac{\varepsilon_1}{g} + \frac{\varepsilon_2}{g^2} + \cdots,$$

are normally distributed on the segment [01].

Proof. If the fractions $\{\alpha g^x\}$ are uniformly distributed, then $P\cdot(1/g^s) + o(P)$ fractions lie on any half interval of the form $\left[\frac{a}{g^s}\ \frac{a+1}{g^s}\right)$. But by Theorem 1, this means that any term $(\delta_1\delta_2\ldots\,\delta_s)$ is encountered in the caterpillar $P\cdot(1/g^s) + o(P)$ times, i.e., the sequence is normal. Conversely, if the sequence is normal, by Theorem 1, $P\cdot(1/g^s) + o(P)$ fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots P$ will fall on any half interval of the type $\left[\frac{a}{g^s}\ \frac{a+1}{g^s}\right)$. In any such half interval δ there will lie $P \operatorname{mes} \delta + O[P(1/g^s)] + o(P)$ fractions (since δ may be approximated with an accuracy of up to $2/g^s$ by a sum of such intervals). Then

$$\overline{\lim_{P\to\infty}}\left|\frac{N_P(\delta)}{P} - \operatorname{mes}\delta\right| = O\left(\frac{1}{g^s}\right).$$

but, letting s approach infinity, we can see that

$$\lim_{P\to\infty}\frac{N_P(\delta)}{P} = \operatorname{mes}\delta,$$

which was to be proved.

Using Champernowne's example of a normal sequence of terms, we construct a number α such that the sequence of fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots$, is uniformly distributed.

For various problems in Diophantine approximations with exponential functions, see Supplement (2).

§4. Criteria for Normality of Base Sequence of Terms

Let us prove the theorem of I.I. Pyatetskiy (see [20]).

Theorem. Let there be a sequence consisting of the terms 0, 1, ...,

g terms of the number α in (2) form a normal sequence of symbols. Conversely, if Sequence (2) is a normal sequence of symbols, then the fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots, P, \ldots$, where

$$\alpha = \frac{\varepsilon_1}{g} + \frac{\varepsilon_2}{g^2} + \cdots,$$

are normally distributed on the segment [01].

Proof. If the fractions $\{\alpha g^x\}$ are uniformly distributed, then $P\cdot(1/g^s) + o(P)$ fractions lie on any half interval of the form $\left[\frac{a}{g^s}\ \frac{a+1}{g^s}\right)$. But by Theorem 1, this means that any term $(\delta_1 \delta_2 \ldots \delta_s)$ is encountered in the caterpillar $P\cdot(1/g^s) + o(P)$ times, i.e., the sequence is normal. Conversely, if the sequence is normal, by Theorem 1, $P\cdot(1/g^s) + o(P)$ fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots P$ will fall on any half interval of the type $\left[\frac{a}{g^s}\ \frac{a+1}{g^s}\right)$. In any such half interval δ there will lie P mes $\delta + O[P(1/g^s)] + o(P)$ fractions (since δ may be approximated with an accuracy of up to $2/g^s$ by a sum of such intervals). Then

$$\overline{\lim_{P\to\infty}}\left|\frac{N_P(\delta)}{P} - \text{mes}\,\delta\right| = O\left(\frac{1}{g^s}\right).$$

but, letting s approach infinity, we can see that

$$\lim_{P\to\infty}\frac{N_P(\delta)}{P} = \text{mes}\,\delta,$$

which was to be proved.

Using Champernowne's example of a normal sequence of terms, we construct a number α such that the sequence of fractions $\{\alpha g^x\}$, $x = 1, 2, \ldots$, is uniformly distributed.

For various problems in Diophantine approximations with exponential functions, see Supplement [2].

§4. Criteria for Normality of Base Sequence of Terms

Let us prove the theorem of I.I. Pyatetskiy (see [20]).

Theorem. Let there be a sequence consisting of the terms 0, 1, ...,

$g-1$,

$$\alpha = a_1 a_2 \ldots \qquad (1)$$

such that there exists a constant C *such that the inequality*

$$\overline{\lim_{X\to\infty}} \frac{N_X(\Delta)}{X} < C\frac{1}{g^s}.$$

will be satisfied for any natural number s *and any* s*-tuple* Δ (or Δ_s) *consisting of the terms* $0, 1, \ldots, g-1$. *Then Sequence (1) is normal*.

We shall follow the proof given by I.I. Pyatetskiy (see [24]). His proof yields a stronger statement: a simpler argument may be used to prove the validity of the criteria.

Lemma 1. *Let us discuss all possible* l*-term combinations of the terms* $0, 1, \ldots, g-1$ (*there are* g^l *of them*). *Let* r *be a natural number. The number of combinations in which we encounter any fixed term a number of times equal to* $l(1/g) + \tau(l/r)$, $|\tau| \geq 1$ *will not exceed* $g^l(r^4/4l^2)$.

Proof. The number of combinations in which the term a is encountered precisely k times will equal

$$C_l^k (g-1)^{l-k},$$

since this sign may be arranged C_l^k ways in k places, while the remaining $l-k$ positions may be filled with any terms except a. The number sought in this lemma equals:

$$\Phi = \sum_{\left|k - l\frac{1}{g}\right| > \frac{l}{r}} C_l^k (g-1)^{l-k} = g^l \sum_{\left|k - l\frac{1}{g}\right| > \frac{l}{r}} C_l^k \left(\frac{1}{g}\right)^k \left(1-\frac{1}{g}\right)^{l-k}.$$

This quantity is found by a well-known method. Since summation is carried out over those k for which $(r^4/l^4)[k - l(1/g)]^4 \geq 1$,

$$\Phi < g^l \frac{r^4}{l^4} \sum_{k=0}^{l} C_l^k \left(k - l\frac{1}{g}\right)^4 \left(\frac{1}{g}\right)^k \left(1-\frac{1}{g}\right)^{l-k} =$$

$$= g^l \frac{r^4}{l^4}\left(s_4 - 4l\frac{1}{g}s_3 + 6l^2\left(\frac{1}{g}\right)^2 s_2 - 4l^3\left(\frac{1}{g}\right)^3 s_1 + l^4\left(\frac{1}{g}\right)^4\right),$$

where

$$s_\tau = \sum_{j=0}^{l} j^\tau C_l^j \left(\frac{1}{g}\right)^j \left(1-\frac{1}{g}\right)^{l-j}, \quad \tau = 1, 2, 3, 4.$$

Let us compute s_τ. We differentiate (with respect to $\underline{x}$) the identity

$$\sum_{j=0}^{l} C_l^j x^j \left(1-\frac{1}{g}\right)^{l-j} = \left(x+1-\frac{1}{g}\right)^l$$

and multiply by $\underline{x}$; we then obtain the identity

$$\sum_{j=0}^{l} jC_l^j x^j \left(1-\frac{1}{g}\right)^{l-j} = lx\left(x+1-\frac{1}{g}\right)^{l-1}.$$

Differentiating this last identity and multiplying by $\underline{x}$, we obtain

$$\sum_{j=0}^{l} j^2C_l^j x^j \left(1-\frac{1}{g}\right)^{l-j} = lx\left(x+1-\frac{1}{g}\right)^{l-1} + l(l-1)x^2\left(x+1-\frac{1}{g}\right)^{l-2}.$$

Repeating these steps

$$\begin{gathered} lx\left(x+1-\frac{1}{g}\right)^{l-1} + 3l(l-1)x^2\left(x+1-\frac{1}{g}\right)^{l-2} + \\ + l(l-1)(l-2)x^3\left(x+1-\frac{1}{g}\right)^{l-3} = \\ = \sum_{j=0}^{l} j^3C_l^j x^j \left(x+1-\frac{1}{g}\right)^{l-j} \end{gathered}$$

and once again repeating this procedure we obtain

$$\begin{gathered} lx\left(x+1-\frac{1}{g}\right)^{l-1} + l(l-1)x^2\left(x+1-\frac{1}{g}\right)^{l-2} + \\ + 6l(l-1)x^2\left(x+1-\frac{1}{g}\right)^{l-2} + \\ + 3l(l-1)(l-2)x^3\left(x+1-\frac{1}{g}\right)^{l-3} + \\ + 3l(l-1)(l-2)x^3\left(x+1-\frac{1}{g}\right)^{l-3} + \\ + l(l-1)(l-2)(l-3)x^4\left(x+1-\frac{1}{g}\right)^{l-4} = \sum_{j=0}^{l} j^4C_l^j x^j \left(x+1-\frac{1}{g}\right)^{l-j}. \end{gathered}$$

Assuming in these formulas that $x = 1/g$, we obtain

$$s_1 = l\frac{1}{g};$$

$$s_2 = l\frac{1}{g} + l(l-1)\frac{1}{g^2};$$

$$s_3 = l\frac{1}{g} + 3l(l-1)\frac{1}{g^2} + l(l-1)(l-2)\frac{1}{g^3};$$

$$\begin{gathered} s_4 = l\frac{1}{g} + 7l(l-1)\frac{1}{g^2} + 6l(l-1)(l-2)\frac{1}{g^3} + \\ + l(l-1)(l-2)(l-3)\frac{1}{g^4}; \end{gathered}$$

$$\Phi < g^l \frac{r^4}{l^4}\Big(l\frac{1}{g} + 7l(l-1)\frac{1}{g^2} + 6l(l-1)(l-2)\frac{1}{g^3} +$$

$$+ l(l-1)(l-2)(l-3)\frac{1}{g^4} -$$

$$-4l^2\frac{1}{g^2} - 12l^2(l-1)\frac{1}{g^3} - 4l^2(l-1)(l-2)\frac{1}{g^4} +$$

$$+ 6l^3\frac{1}{g^3} + 6l^3(l-1)\frac{1}{g^4} -$$

$$-4l^4\frac{1}{g^4} + l^4\frac{1}{g^4}\Big) = g^l\frac{r^4}{l^4}l\cdot\frac{1}{g}\Big(1 - 7\frac{1}{g} + 3l\frac{1}{g} - 6l\frac{1}{g^2} + 12\frac{1}{g^2} +$$

$$+ 3l\frac{1}{g^3} - 6\frac{1}{g^3}\Big) =$$

$$= g^l\frac{r^4}{l^4}l\cdot\frac{1}{g}\Big(1-\frac{1}{g}\Big)\Big(1 + 3(l-2)\frac{1}{g}\Big(1-\frac{1}{g}\Big)\Big).$$

Since $(1/g)[1 - (1/g)] \leq 1/4$,

$$\Phi < g^l\frac{r^4}{l^3}\frac{1}{4}\Big(1 + \frac{3(l-2)}{4}\Big) < g^l\frac{r^4}{4l^2},$$

which has to be proved.

Let the sequence

$$\alpha = a_1, a_2, a_3, \cdots \qquad (1)$$

satisfy the condition of the criterion. We shall first show that the equality

$$\lim_{P\to\infty}\frac{N_P(a)}{P} = \frac{1}{g}.$$

holds for the number of appearances of any term a among the P terms of Sequence (1) (we let $N_P(a)$ stand for this number). We take $l \geq 1$ and combine terms into groups of l components

$$a_1 a_2 \ldots a_l\, a_{l+1} \ldots a_{2l} \ldots$$

We introduce the natural number r. A system of l terms is called "good" if the term a is encountered in it a number of times $l(1/g)$ + + $\theta(l/r)$, $|\theta| \leq 1$; the remaining systems are called "bad." We let $L(P)$ be the number of good systems, and $M(P)$ the number of bad systems up to the $l[P/l]$-th term of α

$$\Big[\frac{P}{l}\Big] = L(P) + M(P).$$

The term a occurs $l(1/g) + \theta(l/r)$ times in a good system, and θl t...

in a bad system. Thus, a term a appears up to the Pth term of α a number of times

$$N_P(a) = L(P)\frac{l}{g} + L(P)\,\theta\,\frac{l}{r} + \theta_1 M(P)\,l + \theta_2 l =$$

$$= \left(\frac{P}{l} - M(P) + \theta_2\right)\frac{l}{g} + \left(\frac{P}{l} - M(P) + \theta_2\right)\theta\,\frac{l}{r} +$$

$$+ \theta_1 M(P)\,l + \theta_2 l =$$

$$= \frac{P}{g} + \frac{P}{r}\,\theta + M(P)\left(\theta_1 l - \frac{l}{g} - \theta\,\frac{l}{r}\right) +$$

$$+ \theta_2\frac{l}{g} + \theta_3\theta\,\frac{l}{r} + \theta_2 l.$$

Here $0 \leq \theta_2 \leq 1$ and $|\theta_3| \leq 1$. The occurrence of a bad system is the occurrence of a bad combination in the caterpillar of rank l. By the hypothesis of the theorem, each bad combination is encountered, when $P \geq P_0$, no more than $2CP\cdot(1/g^l)$ times, and by the lemma there are $O[g^l(r^4/l^2)]$ possible bad systems in all. Consequently

$$M(P) < C_1 P g^l \frac{r^4}{l^2}\cdot\frac{1}{g^l} = C_1 P \frac{r^4}{l^2}.$$

Thus, the term a appears prior to the Pth place a number of times

$$N_P(a) = \frac{P}{g} + \frac{P}{r}\,\theta + O\left(P\frac{r^4}{l} + P\frac{r^4}{l} + P\frac{r^2}{l}\right) + O(l).$$

Hence

$$\overline{\lim_{P\to\infty}}\left|\frac{N_P(a)}{P} - \frac{1}{g}\right| < \frac{1}{r} + O\left(\frac{r^4}{l}\right) + O\left(\frac{r^2}{l}\right).$$

Letting $l \to \infty$ and $r \to \infty$ we find that

$$\lim_{P\to\infty}\frac{N_P(a)}{P} = \frac{1}{g}.$$

We take $s \geq 1$ and any s-termed combination $\Delta = (\delta_1 \ldots \delta_s)$ made up of the symbols 0, 1, ..., $g - 1$. Let us consider the sequences

$$\alpha = a_1 a_2 \ldots \dot{a}_s\, a_{s+1} \ldots \dot{a}_{2s} \ldots$$

$$T\alpha = a_2 a_3 \ldots \dot{a}_{s+1}\, a_{s+2} \ldots \dot{a}_{2s+1} \ldots$$

$$\cdots\cdots\cdots$$

$$T^{s-1}\alpha = a_s a_{s+1} \ldots \dot{a}_{2s-1}\, a_{2s} \ldots \dot{a}_{3s-1} \ldots$$

Each of these sequences may be considered to be a sequence composed

not of g terms but of g^s terms, considering each group between apostrophes $'a_k \dots a_{k+s-1}'$ to be a single term; Δ is such a fixed term. The number of occurrences Δ among the $[P/s]$ terms of the sequence $T^j\alpha$ we call $A_\Delta^{[P/s]}(T^j\alpha)$. Clearly

$$N_P(\Delta) = \sum_{j=0}^{s-1} A_\Delta^{\left[\frac{P}{s}\right]}(T^j\alpha) + O(s).$$

For each of the sequences $T^j\alpha$ the conditions of the theorem will be satisfied. It is clear that the conditions of the theorems will hold for these sequences if they are considered to be sequences composed of g^s terms. Thus

$$\lim_{P\to\infty} \frac{A_\Delta^{\left[\frac{P}{s}\right]}(T^j\alpha)}{\frac{P}{s}} = \frac{1}{g^s}, \quad j = 0, 1, \dots, s-1.$$

Hence

$$\lim_{P\to\infty} \frac{N_P(\Delta)}{P} = \sum_{j=0}^{s-1} \frac{1}{sg^s} = \frac{1}{g^s}.$$

Thus the theorem is proved.

§5. Application of the Criterion of Normality for Sequences of Terms

As A.G. Postnikov has noted [19], using I.I. Pyatetskiy's criterion, it is possible to simplify the proof that the Champernowne sequence is normal.

It has been shown that

$$\lim_{r\to\infty} \frac{N_{X_r}(\Delta)}{X_r} = \frac{1}{g^s}.$$

Let $X_r < X < X_{r+1}$.

$$\frac{N_X(\Delta)}{X} < \frac{N_{X_{r+1}}(\Delta)}{X_r} < \frac{N_{X_{r+1}}(\Delta)}{X_{r+1}} \cdot \frac{X_{r+1}}{X_r}.$$

But

$$\frac{X_{r+1}}{X_r} = \frac{X_r + (r+1)g^{r+1}}{X_r} < 1 + \frac{(r+1)g^{r+1}}{rg^r} < 1 + 2g.$$

Thus

$$\overline{\lim_{X\to\infty}} \frac{N_X(\Delta)}{X} < (1+2g)\frac{1}{g^{s}}.$$

Using the criteria of I.I. Pyatetskiy, we obtain

$$\lim_{X\to\infty} \frac{N_X(\Delta)}{X} = \frac{1}{g^{s}}.$$

This discussion also indicates that it is possible to give an arbitrary order to the r-digit numbers contained in s_r in Champernowne's construction, i.e., it is not necessary to use the natural order as Champernowne did.

§6. A Second Definition of the Normal Sequence of Symbols

We call the infinite sequence consisting of terms 0, 1, ..., $g - 1$ weakly normal (I have deliberately translated Borel's term ([3], page 193) "simplement normal" as weakly normal) if the asymptotic frequency of occurrence of each of the terms 0, 1, ..., $g - 1 = 1/g$.

The sequence made up of the symbols 0, 1, ..., $g - 1$,

$$a_1, a_2, a_3, \ldots \tag{1}$$

is called weakly normal to the scale g^k (k is a fixed natural number) if when the terms of the sequence are combined into groups of k members

$$(a_1a_2\ldots a_k)(a_{k+1}a_{k+2}\ldots a_{2k})\ldots \tag{2}$$

and each parenthesized term is considered to be a symbol in the alphabet consisting of g^k elements, we obtain a weakly normal sequence.

Definition. Sequence (1) is called a normal sequence of symbols if for any natural number k the sequence is weakly normal to the scale g^k.

Theorem. The definitions of normal sequences given in §§ 1 and 6 are equivalent.

This theorem is due to Pillai [30, 31]. A somewhat less complete result is given by Niven and Zuckerman [32] (see [33] as well). There

is a proof of this theorem in a paper by Maxfield [34]; I did not understand it, and have given another proof.

We shall prove that a sequence that is normal in the sense of the definition of this paragraph is normal in the sense of §1.

Let Δ be a fixed k-term combination of the symbols 0, 1, ..., g − 1. Let $A \geq k$ be an increasing natural number. We use all possible A-tuples consisting of the terms 0, 1, ..., g − 1.

$$(b_1 b_2 \ldots b_A).$$

We assign the "measure" $1/g^A$ to each such object. To some set Σ consisting of the various A-tuples we assign the "measure" $\mu\Sigma$, equal to the sum of the measures of the objects entering into the set.

We write each A-tuple as follows:

$$(b_1 \ldots b_k)(b_2 \ldots b_{k+1}) \ldots (b_{A-k+1} \ldots b_A)$$

and calculate how many times the element Δ will be found in this sequence. Let Σ_v be the set of such elements in which Δ is encountered v times.

Lemma (Markov). When $A \to \infty$

$$1 \cdot \mu\Sigma_1 + 2 \cdot \mu\Sigma_2 + \ldots + (A-k+1) \cdot \mu\Sigma_{A-k+1} = \frac{1}{g^k} A + o(A).$$

The proof may be found in [35], Chapter 6.

We let $T_l(\Sigma)$ be the number of occurrences of some set of A-tuples Σ prior to the lth term of the sequence

$$(a_1 \ldots a_A)(a_{A+1} \ldots a_{2A}) \ldots$$

Let $P = Al + r$, $0 \leq r \leq A - 1$,

$$N_P(\Delta) = T_l(\Sigma_1) + 2T_l(\Sigma_2) + \ldots + (A-k+1)\, T_l(\Sigma_{A-k+1}) + O(lk) + O(r).$$

From this and from the hypothesis of the theorem it follows that

$$\overline{\lim_{P\to\infty}} \left| \frac{N_P(\Delta)}{P} - \frac{1}{g^k} \right| = O\left(\frac{k}{A}\right).$$

Letting A approach ∞, we obtain $\lim\limits_{P\to\infty} \frac{N_P(\Delta)}{P} = \frac{1}{g^k}$, which was what we were to prove.

Conversely, let there be given a sequence normal in the sense of §1

$$a_1, a_2, \ldots, a_p, \ldots \tag{3}$$

We combine Sequence (3) into terms of k elements each

$$A_1, A_2, \ldots, \tag{4}$$

where $A_l = (a_{1+(l-1)k}, \ldots, a_{k-1+(l-1)k})$.

We shall prove that Sequence (4) is normal. We let $\omega = (B_1 \ldots B_s)$, where $B_i = (\delta_1^{(i)} \ldots \delta_k^{(i)})$, $i = 1, 2, \ldots, s$, while the $\delta_j^{(i)}$ are taken from the terms $0, 1, \ldots, g-1$.

We let $\tilde{N}_p(\omega)$ be the number of appearances of the term ω prior to the Pth term of the sequence

$$(A_1 \ldots A_s)(A_2 \ldots A_{s+1}) \ldots \tag{5}$$

We let $\Omega = (\delta_1^{(1)} \ldots \delta_k^{(1)} \ldots \delta_1^{(s)} \ldots \delta_k^{(s)})$.

It is clear that

$$\tilde{N}_P(\omega) < N_{Pk}(\Omega).$$

Hence

$$\overline{\lim_{P\to\infty}} \frac{N_P(\Delta)}{P} < k \frac{1}{g^{ks}}.$$

By the criterion of §4, the sequence is normal. This means that it is weakly normal. The fact that k is arbitrary means that Sequence (3) satisfies the definition of normality given in this section.

Let α be an absolutely normal Borel number (see the Introduction). Let the base-g expansion be as follows:

$$\alpha = \frac{a_1}{g} + \frac{a_2}{g^2} + \ldots$$

The base-g^k decomposition of α will be

$$\alpha = \frac{a_1 g^{k-1} + \ldots + a_k}{g^k} + \frac{a_{k+1} g^{k-1} + \ldots + a_{2k}}{g^{2k}} + \ldots$$

By the condition for weak normality for the base g^k any combination $(b_1 \ldots b_k)$ will be contained in the sequence

$$(a_1 \ldots a_k)(a_{k+1} \ldots a_{2k}) \ldots$$

with an asymptotic frequency $1/g^k$. This means that the sequence $a_1a_2\ldots$ is a normal sequence of terms.

§7. A System of Mutually Normal Sequences of Terms

We introduce here a concept due to N.M. Korobov ([36], page 363).

We are given integers $g_1, g_2, \ldots, g_k$ larger than or equal to two. We are given the system of infinite sequences of terms

$$\begin{aligned} \alpha_1 &= a_{11}\, a_{12} \ldots \\ \alpha_2 &= a_{21}\, a_{22} \ldots \\ &\ldots\ldots\ldots \\ \alpha_k &= a_{k1}\, a_{k2} \ldots, \end{aligned} \qquad (1)$$

where the sequence α_j consists of the symbols $0,1, \ldots, g_j - 1 (j = 1, \ldots, k)$.

Let $s \geq 1$. The sequence of matrixes

$$\begin{pmatrix} a_{11}\ldots a_{1s} \\ a_{21}\ldots a_{2s} \\ \ldots\ldots \\ a_{k1}\ldots a_{ks} \end{pmatrix} \begin{pmatrix} a_{12}\ldots a_{1s+1} \\ a_{22}\ldots a_{2s+1} \\ \ldots\ldots \\ a_{k2}\ldots a_{ks+1} \end{pmatrix} \ldots \begin{pmatrix} a_{1P}\ldots a_{1P+s-1} \\ a_{2P}\ldots a_{2P+s-1} \\ \ldots\ldots \\ a_{kP}\ldots a_{kP+s-1} \end{pmatrix} \qquad (2)$$

we call a caterpillar of length P (rank $\underline{s}$) of Sequence (1).

We take any matrix

$$\Delta_s = \begin{pmatrix} b_{11}\ldots b_{1s} \\ b_{21}\ldots b_{2s} \\ \ldots\ldots \\ b_{k1}\ldots b_{ks} \end{pmatrix},$$

in which the $\underline{jth}$ row is composed of the symbols $0, 1, \ldots, g_j - 1$. The number of appearances of Δ_s in the caterpillar of length P of the system of sequences (1) we call $N_P(\Delta_s)$. We call the system of sequences (1) mutually normal if for any natural number $\underline{s}$ and any matrix Δ_s

$$\lim_{P\to\infty} \frac{N_P(\Delta_s)}{P} = \frac{1}{(g_1 g_2 \ldots g_k)^s}.$$

The theory of mutually normal sequences of terms reduces to a theory of normal sequences of terms.

We shall consider any column $\begin{pmatrix} b_1 \\ \vdots \\ b_k \end{pmatrix}$ to be a single symbol, where $0 \leq b_j \leq g_j - 1$, $j = 1, 2, \ldots k$. It is clear that $g_1, g_2, \ldots, g$

such terms. The system of sequences (1) may be considered to be a single sequence

$$\bar{a} = \bar{a}_1, \bar{a}_2, \bar{a}_3, \ldots, \tag{1'}$$

but where the terms are taken from an alphabet containing $g_1 g_2 \cdots g_k$ symbols. It is clear that the mutual normality of the system of sequences (1) is equivalent to the normality (in the normal sense) of the auxiliary sequence (1'). From this there follows a theorem.

Theorem. <u>Let the system of sequences (1) be such that there exists a constant</u> C <u>such that</u>

$$\lim_{P \to \infty} \frac{N_P(\Delta_s)}{P} < C \frac{1}{(g_1 \cdots g_k)^s}.$$

<u>for any natural number</u> s <u>taken or for any matrix</u> Δ_s <u>considered. Then the system of sequences is mutually normal</u>.

This is a possible method for constructing a system of mutually normal sequences. The normal sequences constructed of $g_1 \cdots g_{k-1}$ terms. Each term $0 \leq a \leq g_1 \cdots g_{k-1}$ corresponds to a column of $\underline{k}$ rows (the $\underline{j\text{th}}$ row consists of the numbers 0, 1, ..., g_{j-1}) and the initial row is expanded into $\underline{k}$ rows by this correspondence.

Example. Let $g_1 = 2$, $g_2 = 3$. We establish the correspondence

$$0 \to \binom{0}{0}, 1 \to \binom{1}{0}, 2 \to \binom{0}{1}, 3 \to \binom{1}{1}, 4 \to \binom{0}{2}, 5 \to \binom{1}{2}.$$

We use a Champernowne row for six symbols

0'1'2'3'4'5'00'01'02'03'04'05'10'11'12'13'14'15'20'21'22'23'24'25'30'31'32
33'34'35'40'41'42'43'44'45'50'51'52'53'54'55'...

It corresponds to the system of sequences

0'1'0'1'0'1'00'01'00'01'00'01'10'11'10'11'10'11'
0 0 1 1 2 2 00 00 01 01 02 02 00 00 01 01 02 02

00'01'00'01'00'01'10'11'10'11'10'11'00'01'00'01'
10 10 11 11 12 12 10 10 11 11 12 12 20 20 21 21

00'01'10'11'10'11'10' ...
22 22 20 20 21 21 22 ...

The system of sequences obtained is mutually normal.

§8. The Problem of Constructing a Sequence Mutually Normal to a Given Sequence

The material presented here has been taken from L.P. Starchenko [37, 38].

For a natural number $g \geq 2$, let us consider a normal sequence consisting of the symbols 0, 1, ..., $g - 1$,

$$a_1 a_2 \ldots \tag{1}$$

It is necessary to construct a sequence of terms

$$\delta_1 \delta_2 \ldots, \tag{2}$$

such that the system of sequences (3)

$$\begin{matrix} a_1 a_2 \ldots \\ \delta_1 \delta_2 \ldots \end{matrix} \tag{3}$$

will be mutually normal.

We shall first prove that for any natural number <u>s</u> there is an $\Phi(s)$ such that for any s-term combination Δ_s

$$\frac{N_{P'}(\Delta_s) - N_{P''}(\Delta_s)}{P' - P''} < \frac{C}{g^s}.$$

provided that $P'' > \Phi(s)$ and $P' - P'' > P''$

Actually

$$\frac{N_{P'}(\Delta_s) - N_{P''}(\Delta_s)}{P' - P''} = \frac{\frac{N_{P'}(\Delta_s)}{P'} - \frac{N_{P''}(\Delta_s)}{P'}}{1 - \frac{P''}{P'}} =$$

$$= \frac{\frac{N_{P'}(\Delta_s)}{P'} - \frac{N_{P''}(\Delta_s)}{P''}}{1 - \frac{P''}{P'}} + \frac{\frac{N_{P''}(\Delta_s)}{P''} - \frac{N_{P''}(\Delta_s)}{P'}}{1 - \frac{P''}{P'}}.$$

According to the Cauchy convergence criterion there exists a $P(s, \Delta_s)$ such that for every P' and P'' larger than $P(s, \Delta_s)$ the inequality

$$\left| \frac{N_{P'}(\Delta_s)}{P'} - \frac{N_{P''}(\Delta_s)}{P''} \right| < \frac{1}{g^s}.$$

holds. By the normality of the initial sequence we have where $P'' > \tilde{P}(\Delta_s)$

$$\frac{N_{P''}(\Delta_s)}{P''} < \frac{C_1}{g^s}.$$

We take the largest of the numbers $\tilde{P}(\Delta_s)$ and $P(s, \Delta_s)$ taken over all possible Δ_s as $\Phi(s)$.

Taking these inequalities into account, we obtain

$$\frac{N_{P'}(\Delta_s) - N_{P''}(\Delta_s)}{P' - P''} < \frac{1}{g^s}\frac{P'}{P' - P''} + N_{P''}(\Delta_s)\frac{\frac{1}{P''} - \frac{1}{P'}}{1 - \frac{P''}{P'}} =$$

$$= \frac{1}{g^s}\left(\frac{P''}{P' - P''} + 1\right) + \frac{N_{P''}(\Delta_s)}{P''} < \frac{1}{g^s}\left(1 + \frac{P''}{P' - P''}\right) + \frac{C_1}{g^s} < \frac{C}{g^s}.$$

We now proceed to the construction: to do this we will "distort" the initial sequence to an ever greater degree.

In the first line, we write the given normal sequence (1), and in the second a sequence that we can prove is mutually normal with respect to the given sequence.

$$\left.\begin{array}{l} '\varepsilon_1\varepsilon_2 \ldots\ldots \varepsilon_{\Phi(2)-1} \\ \varepsilon_1\varepsilon_2 \ldots\ldots \varepsilon_{\Phi(2)-1} \end{array}\right\} \text{0-th series}$$

$$\left.\begin{array}{l} '\varepsilon_{\Phi(2)} \ldots \varepsilon_{2\Phi(2)}\ '\varepsilon_{2\Phi(2)+1} \ldots \varepsilon_{3\Phi(2)+2}\ '\varepsilon_{3\Phi(2)+3} \ldots \varepsilon_{\overline{\Phi(4)}-1} \\ \varepsilon_{\Phi(2)+1} \ldots \varepsilon_{2\Phi(2)+1}\ '\varepsilon_{2\Phi(2)+3} \ldots \varepsilon_{3\Phi(2)+3}\ '\varepsilon_{3\Phi(2)+4} \ldots \varepsilon_{\overline{\Phi(4)}} \end{array}\right\} \text{1-st series}$$

$$\left.\begin{array}{l} '\varepsilon_{\overline{\Phi(4)}} \ldots \varepsilon_{2\overline{\Phi(4)}}\ '\varepsilon_{2\overline{\Phi(4)}+1} \ldots \varepsilon_{3\overline{\Phi(4)}+2}\ '\varepsilon_{3\overline{\Phi(4)}+3} \ldots \varepsilon_{\overline{\Phi(6)}-1} \\ \varepsilon_{\overline{\Phi(4)}+2} \ldots \varepsilon_{2\overline{\Phi(4)}+2}\ '\varepsilon_{2\overline{\Phi(4)}+3} \ldots \varepsilon_{3\overline{\Phi(4)}+4}\ '\varepsilon_{3\overline{\Phi(4)}+5} \ldots \varepsilon_{\overline{\Phi(6)}+1} \end{array}\right\} \text{2-nd series}$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

where $\overline{\Phi(2)} = \Phi(2)$,

$$\overline{\Phi(2s)} = 2^{v_s}\,\overline{\Phi(2s-2)} + 2^{v_s} - 1 \quad (s = 2, 3, \ldots),$$

while v_s is any natural number for which

$$\Phi(2s) < 2^{v_s}\,\overline{\Phi(2s-2)} + 2^{v_s} - 2.$$

Then

$$\overline{\Phi(2s)} > \Phi(2s).$$

We shall prove that we have two mutually normal sequences.

We take any k-column matrix (k is fixed)

$$\bar{\Delta}_k = \binom{\Delta_k}{\Delta'_k}.$$

Let

$$X = 2^{v}\,\overline{\Phi(2s)} + 2^{v} - 2,$$

where $v = 0, 1, \ldots, v_{s+1} - 1$.

We let G_X be the number of times $\bar{\Delta}_k$ appears in the caterpillar

of rank $\underline{k}$ of length X for the described system of sequences.

Let $s \geq k$.

Prior to the number $\overline{\Phi(2k)} - 1$, the combination $\overline{\Delta}_k$ is encountered some number of times. We let this number be L. The appearance of $\overline{\Delta}_k$ in an s-series indicates that in the corresponding positions of the caterpillar of rank 2s of the initial sequence there appears one of the $g^{2(s-k)}$ combinations having the form

$$(\underbrace{\Delta_k \ldots}_{s\text{ terms}} \underbrace{\Delta'_k \ldots}_{s\text{ terms}}).$$

(This is valid with the possible exception of the last 2s numbers of the series.) In each section separated by apostrophes, such a combination appears less than $(C/g^{2s})Q$ times, where Q is the number of symbols in this section; but the number of such combinations equals $g^{2(s-k)}$, and thus the number of occurrences of $\overline{\Delta}_k$ will be less than $g^{2(s-k)}CQ/g^{2s} = CQ/g^{2k}$. Thus

$$G_X < L + \frac{CX}{g^{2k}} + O(s^2).$$

The quantity $O(s^2)$ appears at series splices.

$$\frac{G_X}{X} < \frac{L}{X} + \frac{C}{g^{2k}} + O\left(\frac{s^2}{X}\right).$$

$s^2/X \to 0$, since $X > C2^s$ where C is some constant.

Thus

$$\overline{\lim_{X\to\infty}} \frac{G_X}{X} < \frac{C}{g^{2k}}.$$

Let

$$X_n < P < X_{n+1},$$

$$\text{where } X_{n+1} = 2X_n + 1.$$

$$\frac{N_P(\overline{\Delta}_k)}{P} < \frac{N_{X_{n+1}}(\overline{\Delta}_k)}{X_{n+1}} = \frac{N_{X_{n+1}}(\overline{\Delta}_k)}{X_{n+1}} \cdot \frac{X_{n+1}}{X_n},$$

$$\frac{X_{n+1}}{X_n} \to 2 \text{ when } P \to \infty.$$

$$\overline{\lim_{P\to\infty}} \frac{N_P(\overline{\Delta}_k)}{P} < \frac{2C}{g^{2k}}.$$

Since the criteria for mutual normality of the two series have been satisfied, the required system of sequences has been constructed.

Let us construct an infinite number of series mutually normal to the given Sequence (1). We construct one sequence mutually normal to Sequence (1)

$$\begin{matrix} \varepsilon_1 \varepsilon_2 \ldots \varepsilon_s \ldots \\ \delta_1 \delta_2 \ldots \delta_s \ldots \end{matrix} \tag{3}$$

Sequence (3) is a normal sequence in which the terms are taken from an alphabet consisting of g^2 elements (a column is considered to be a term). We construct a sequence mutually normal to Sequence (3) and consisting of g^2 symbols

$$\begin{matrix} \varepsilon_1 \varepsilon_2 \varepsilon_3 \ldots \varepsilon_s \ldots \\ \delta_1 \delta_2 \delta_3 \ldots \delta_s \ldots \\ \varepsilon'_1 \varepsilon'_2 \varepsilon'_3 \ldots \varepsilon'_s \ldots \\ \delta'_1 \delta'_2 \delta'_3 \ldots \delta'_s \ldots \end{matrix} \tag{4}$$

Sequence (4) may be considered to be a normal sequence in which the symbols are taken from an alphabet consisting of g^4 elements. We construct a sequence mutually normal to Sequence (4), etc.

§9. Bernoulli-Normal Sequences of Terms

Let there be given two positive numbers p and q such that $p + q = 1$. Let there be an infinite sequence consisting of symbols 0 and 1,

$$\varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots \tag{1}$$

Let s be any natural number. We write Sequence (1) in the form of a "caterpillar"

$$(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_s)(\varepsilon_2 \varepsilon_3 \ldots \varepsilon_{s+1}) \ldots (\varepsilon_p \varepsilon_{p+1} \ldots \varepsilon_{p+s-1}) \ldots \tag{2}$$

Let $\Delta = (\delta_1 \ldots \delta_s)$ be any s-tuple consisting of the symbols 0 and 1. We let $N_p(\Delta)$ be the number of occurrences of the term Δ prior to the pth term of Sequence (2). We call Sequence (1) a Bernoulli-normal sequence of symbols if for any natural s and any s-tuple $\Delta = (\delta_1 \ldots \delta_s)$,

$$\lim_{P\to\infty}\frac{N_P(\Delta)}{P}=p^j q^{s-j},$$

where j is the number of ones among $(\delta_1 \ldots \delta_s)$.

Let us prove the following generalization of I.I. Pyatetskiy's criterion (see [12]).

Theorem. Let the sequence of symbols 1 and 0

$$\alpha = \varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots$$

be such that there exists a constant $C > 0$ such that for any natural number s and any s-tuple $\Delta = (\delta_1 \ldots \delta_s)$ consisting of 0 and 1

$$\lim_{P\to\infty}\frac{N_P(\Delta)}{P}< C p^j q^{s-j},$$

where j is the number of ones among the $\delta_1\delta_2 \ldots \delta_s$.

Then Sequence (1) is a Bernoulli-normal sequence of terms.

The proof is similar to the proof of the theorem in §4.

If $\Delta = (\delta_1 \ldots \delta_s)$ is some element consisting of zeroes and ones, we then let $\mu\Delta = p^j q^{s-j}$ where j is the number of ones among the $\delta_1 \ldots \delta_s$. We will call the quantity $\mu\Delta$ the measure of the element. If Σ is some set of different s-tuples we assume that $\mu\Sigma$ equals the sum of the measures of the multiple-term elements entering into the set.

Lemma 1. Let r be a natural number. Let Σ be the set of those l-tuples in which the number of ones j satisfies the inequality $|j - lp| \geq l/r$.

Then

$$\mu\Sigma < \frac{r^2}{4l}.$$

As in Lemma 1 of §4 we find that $\Phi = \mu\Sigma$ and equals

$$\Phi = \sum_{\substack{k=0\\ |k-lp|\geq \frac{l}{r}}}^{l} C_l^k p^k q^{l-k}.$$

Carrying through the procedure used in Lemma 1 of §4, we obtain the evaluation needed.

Lemma 2. Let r be a natural number, Δ some s-termed combination. Let l be a natural number.

Let us consider the ls-tuple consisting of symbols 0 and 1

$$(a_1 a_2 \dots a_s a_{s+1} \dots a_{2s} \dots a_{(l-1)s+1} \dots a_{ls})$$

and let us group it into s elements such that we may use the representation

$$(b_1 b_2 \dots b_s),$$

where $b_k = (a_{(k-1)s+1} \dots a_{ks})$ (for s = 1, both forms are the same). We let $A_\Delta^{(l)}$ be the number of occurrences of the combination Δ among the $b_1 b_2 \dots b_s$. We let $\mathfrak{M}$ be the set of those ls-tuples for which $|A_\Delta^l - l\mu\Delta| \geq$ $\geq l/r$ (r is any natural number). Then

$$\mu\mathfrak{M} < \frac{r^4}{4l^2}.$$

Proof. The measure sought equals

$$\sum_{\substack{k=0 \\ |k - l\mu\Delta| > \frac{l}{r}}}^{l} C_l^k (\mu\Delta)^k (1 - \mu\Delta)^{l-k}.$$

Repeating the argument used to prove Lemma 1 of §4, we obtain the statement required.

Let us now prove the theorem. Let the conditions of the criterion be satisfied for Sequence (1). We group the terms of Sequence (1) into s elements

$$b_1, b_2, \dots, b_n, \dots, \tag{3}$$

where $b_n = (a_{(n-1)s+1} \dots a_{ns})$. From the P terms of Sequence (1) it is possible to form [P/s] terms of Sequence (3). Let us take a natural number l and group the terms of Sequence (3) into elements of l components

$$\underbrace{b_1 b_2 \dots b_l}\ \underbrace{b_{l+1} \dots b_{2l}} \dots \tag{3}$$

We select a natural number r, and will call an l-term group good if Δ occurs in it a number of times $l[\mu\Delta + \theta(1/r)]$, $|\theta| \leq 1$; we shall call it bad if this is not the case. We let $N([P/s])$ be a

number indicating how many times a bad group is found among the first $\underline{l}[[P/s]/\underline{l}]$ terms of Sequence (3). Then a good group occurs a number of times

$$\frac{P}{sl} - M\left(\left[\frac{P}{s}\right]\right) + O(1)$$

with an absolute constant in O. A good group introduces $l[\mu\Delta + (\theta/r)]$ symbols, while a bad group yields no more than $\underline{l}$. Thus the number of times the term Δ occurs prior to the $[P/s]$ position in Sequence (3) is

$$A_{\Delta}^{\left[\frac{P}{s}\right]}(s) = l\left(\mu\Delta + \frac{\theta}{r}\right)\left(\frac{P}{sl} - M\left(\left[\frac{P}{s}\right]\right) + O(1)\right) + l\theta_1 M\left(\left[\frac{P}{s}\right]\right) + O(l),$$

where $0 \leq \theta_1 \leq 1$, $O(\underline{l})$ occurs owing to the fact that there possibly is an incomplete group and thus there is an absolute constant in O. Each bad combination belongs to a system $\mathfrak{M}$; thus

$$M\left(\left[\frac{P}{s}\right]\right) < N_{s\left[\frac{P}{s}\right]}(\mathfrak{M}).$$

But when $P \geq P_0$ by hypothesis

$$N_{s\left[\frac{P}{s}\right]}(\mathfrak{M}) < 2C\cdot\mu\mathfrak{M}\cdot P.$$

Moreover, by Lemma 2, $\mu\mathfrak{M} \leq r^4/4\underline{l}^2$. Consequently when $P \geq P_0$

$$A_{\Delta}^{\left[\frac{P}{s}\right]}(s) = \frac{P}{s}\left(\mu\Delta + \frac{\theta}{r}\right) + O\left(P\frac{r^4}{l}\right) + O(l).$$

$$\overline{\lim_{P\to\infty}}\left|\frac{A_{\Delta}^{\left[\frac{P}{s}\right]}(s)}{\frac{P}{s}} - \mu\Delta\right| < \frac{1}{r} + O\left(\frac{sr^4}{l}\right).$$

When $\underline{l}$ approaches infinity we obtain

$$\overline{\lim_{P\to\infty}}\left|\frac{A_{\Delta}^{\left[\frac{P}{s}\right]}(s)}{\frac{P}{s}} - \mu\Delta\right| < \frac{1}{r}.$$

When $\underline{r}$ approaches infinity we obtain

$$\lim_{P\to\infty} \frac{A_\Delta^{\left[\frac{P}{s}\right]}(\alpha)}{\frac{P}{s}} = \mu\Delta .$$

We let $T^j\alpha$, $j = 1, 2, \ldots, s - 1$ be the sequences

$$T^j\alpha = \varepsilon_{1+j}, \varepsilon_{s+j}, \varepsilon_{2s+j}, \ldots \tag{4}$$

We have

$$\lim_{P\to\infty} \frac{A_\Delta^{\left[\frac{P}{s}\right]}(T^j\alpha)}{\frac{P}{s}} = \mu\Delta,\ j = 0, 1, \ldots, s - 1 .$$

But it is clear that

$$N_P(\Delta) = \sum_{j=0}^{s-1} A_\Delta^{\left[\frac{P}{s}\right]}(T^j\alpha) + O(s).$$

And from this it follows that

$$\lim_{P\to\infty} \frac{N_P(\Delta)}{P} = s\frac{\mu\Delta}{s} = \mu\Delta,$$

which was what we were to prove.

We need a concept that is more general than the concept of a Bernoulli-normal sequence of terms.

Let there be given g positive numbers $p_0, \ldots p_{g-1}$; let $p_0 + p_1 + \ldots + p_{g-1} = 1$.

Let us consider an s-tuple made up of symbols $0, 1, \ldots, g - 1$, $\Delta = (\delta_1 \ldots \delta_s)$.

We let $\mu\Delta$ be the quantity $p_{\delta 1} \ldots p_{\delta s}$. It is clear that

$$\sum_{\Delta_s} \mu\Delta_s = (p_0 + \ldots + p_{g-1})^s = 1,$$

where the sum in the left side extends over all s-tuples.

We call the sequence composed of the terms $0, 1, \ldots, g - 1$,

$$\alpha_1, \alpha_2, \ldots, \alpha_P \ldots \tag{5}$$

a normal realization of independent trials if for any s and any combination Δ the relationship

$$\lim_{P\to\infty} \frac{N_P(\Delta)}{P} = \mu\Delta .$$

holds.

The criteria are valid.

Theorem. Let Sequence (5) be such that there exists a constant C such that

$$\overline{\lim_{P\to\infty}} \frac{N_P(\Delta)}{P} < C\mu\Delta$$

for any combination of any length. Then Sequence (5) is a normal realization of independent trials.

The proof generalizes in obvious manner the proof of the criterion for Bernoulli-normal sequences of symbols.

§10. Construction of a Bernoulli-Normal Sequence of Symbols

Here we present the construction of a Bernoulli-normal sequence of symbols given by A.G. Postnikov and I.I. Pyatetskiy (see [12]).

We use an idea due to Champernowne [18].

Let p (as before) be the probability for the occurrence of an event in each trial. We take any sequence of rational numbers α_r/β_r such that

$$p = \lim \frac{\alpha_r}{\beta_r}, \quad 0 < \alpha_r < \beta_r,$$

$$\beta_1 < \beta_2 \leq \beta_3 \dots \frac{\beta_r}{\beta_{r-1}} = 1 + O\left(\frac{1}{r}\right).$$

Such a sequence exists for any p, $0 < p < 1$; if p is a rational number, $p = \alpha/\beta$, we then simply take $\alpha_r = \alpha$, $\beta_r = \beta$ for any r.

We let s_r be the sequence of all r-digit numbers in the binary system; a number in which a one is encountered ν times and a zero, consequently, $r - \nu$ times will be repeated $\alpha_r^{\nu}(\beta_r - \alpha_r)^{r-\nu}$ times. We will separate the numbers by apostrophes. For example let $\alpha_3 = 2$, $\beta_3 = 3$. Then

$$s_3 = 000'001'001'010'010'011'011'011'011'100'100'101'101'101'101'110'110'110'110'111'111'111'111'111'111'111'111'.$$

Let us prove that the sequence of terms s_r, written symbolically as

$$\alpha = s_1, s_2, s_3, \ldots$$

is Bernoulli-normal. To do this we must show that any s-termed combination in which there are ν ones is encountered with an asymptotic frequency $p^{\nu}q^{s-\nu}$. In accordance with the criterion it is sufficient to show that there exists an absolute constant C (independent of Δ) such that

$$\overline{\lim_{P \to \infty}} \frac{N_P(\Delta_s)}{P} < Cp^{\nu}q^{s-\nu}.$$

We let x_r be the number of terms in s_r, S_r the sequence $s_1 s_2 \ldots s_r$, X_r the number of terms in $S_r (X_r = x_1 + x_2 + \ldots + x_r)$, g_r the number of occurrences of Δ_s in s_r, and G_r the number of occurrences of Δ_s in S_r. Let us calculate x_r. The number of r-digit numbers in which a one is encountered k times will equal C_r^k; each number will be repeated $\alpha_r^{\nu}(\beta_r - \alpha_r)^{r-\nu}$ times. Thus, there will be

$$\sum_{k=0}^{r} C_r^k \alpha_r^k (\beta_r - \alpha_r)^{r-k} = \beta_r^r,$$

r-digit numbers in s_r, and $x_r = r\beta_r^r$ terms.

In s_r the Δ_s may or may not be separated by an apostrophe. If $r < s$ then Δ_s cannot be contained undivided in s_r. If $r \geq s$, Δ_s is contained undivided in s_r exactly

$$(r - s + 1)\, \alpha_r^{\nu} (\beta_r - \alpha_r)^{s-\nu} \beta_r^{r-s}$$

times. Actually, there exist $r - s + 1$ ways in which Δ_s can occupy an undivided position in an r-digit number (the first symbol of Δ_s may coincide with the first term, with the second term, ..., with the $r - s + 1$-th term of the r-digit number). Δ_s occupies s places in this number, and in the remaining $r - s$ places we may place f ones ($0 \leq f \leq r - s$) in C_{r-s}^f ways, while the remaining places are filled with zeroes. Such an r-digit number must repeat $\alpha_r^{\nu+f}(\beta_r - \alpha_r)^{r-\nu-f}$ times (to it we add ν ones from Δ_s and f). Thus, Δ_s is contained undivided in s_r exactly

$$(r-s+1)\sum_{l=0}^{r-s} C_{r-s}^{l}\alpha_r^{\nu+l}(\beta_r-\alpha_r)^{r-\nu-l} =$$

$$=(r-s+1)\alpha_r^{\nu}(\beta_r-\alpha_r)^{s-\nu}\sum_{l=0}^{r-s} C_{r-s}^{l}\alpha_r^{l}(\beta_r-\alpha_r)^{r-s-l} =$$

$$=(r-s+1)\alpha_r^{\nu}(\beta_r-\alpha_r)^{s-\nu}\beta_r^{r-s}$$

times. s_r contains β_r^r apostrophes. A given apostrophe cannot separate more than s different Δ_s. Thus Δ_s is contained divided no more than $O(\beta_r^r)$ times (s is introduced into the symbol O since s does not increase).

Thus,

$$g_r=(r-s+1)\alpha_r^{\nu}(\beta_r-\alpha_r)^{s-\nu}\beta_r^{r-s}+O(\beta_r^r)=$$

$$=r\beta_r^r\left(\frac{\alpha_r}{\beta_r}\right)^{\nu}\left(1-\frac{\alpha_r}{\beta_r}\right)^{s-\nu}+O(\beta_r^r)=x_r\left(\frac{\alpha_r}{\beta_r}\right)^{\nu}\left(1-\frac{\alpha_r}{\beta_r}\right)^{s-\nu}+o(x_r).$$

Since

$$\left(\frac{\alpha_r}{\beta_r}\right)^{\nu}\left(1-\frac{\alpha_r}{\beta_r}\right)^{s-\nu}=p^{\nu}q^{s-\nu}+o(1).$$

Then

$$g_r=x_r p^{\nu} q^{s-\nu}+o(x_r).$$

Further

$$G_r=\sum_{k=1}^{r} g_k+O(r), \quad X_r=\sum_{k=1}^{r} x_r.$$

We obtain

$$\lim_{r\to\infty}\frac{G_r}{X_r}=p^{\nu}q^{s-\nu}.$$

Let

$$X_{r-1}\leqslant P<X_r,$$

then $N_P(\Delta_s) \leq G_r$ and $(r-1)\beta_{r-1}^{r-1} \leq X_{r-1}$.

Moreover, $X_r = X_{r-1} + r\beta_r^r$. From this it follows that

$$\frac{1}{P}<\frac{X_r}{X_{r-1}}\cdot\frac{1}{X_r}=\left(1+\frac{r\beta_r^r}{X_{r-1}}\right)\frac{1}{X_r}<\left(1+\frac{r\beta_r^r}{(r-1)\beta_{r-1}^{r-1}}\right)\frac{1}{X_r}$$

while since $\beta_r/\beta_{r-1} = 1 + O(1/r)$,

$$\frac{1}{P} < C\frac{1}{X_r},$$

$$\overline{\lim_{P\to\infty}} \frac{X_P(\Delta_s)}{P} < C \lim_{r\to\infty} \frac{G_r}{X_r} = Cp^{\nu} q^{s-\nu}.$$

The criteria are satisfied and our statement is proved.

§11. Relationship of the Concepts of Bernoulli-Normal Sequence of Symbols and Admissible Number

A. Copeland [14] has introduced the notion of an admissible number. Let there be two positive numbers p and q such that $p + q = 1$. Consider the infinite sequence formed of the symbols 0 and 1,

$$\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots. \qquad (1)$$

This sequence is called an admissible number if for any natural number m and for any different nonnegative integers $r_1, r_2, \ldots, r_k$ less than m the sequence

$$\beta_0, \beta_1, \beta_2, \ldots \qquad (2)$$

where $\beta_n = \varepsilon_{nm+r_1}\varepsilon_{nm+r_2}\cdots\varepsilon_{nm+r_k}$ ($n = 1, 2, \ldots$) possesses the property that the relative frequency with which a one appears in Sequence (2) approaches p^k when $n \to \infty$.

Let us prove a theorem.

Theorem. The notion of a Bernoulli-normal sequence of symbols is coextensive with the notion of the Copeland admissible number.

Proof. Consider the Bernoulli-normal sequence of symbols

$$\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots \qquad (1)$$

We take arbitrary m and combine the symbols in Sequence (1) into groups of m components

$$a_1, a_2, a_3, \ldots, \qquad (2)$$

where $a_t = (\varepsilon_{mt}, \ldots, \varepsilon_{mt+m-1})$, $t = 0, 1, \ldots$. The terms of Sequence (2) are taken from an alphabet containing 2^m elements.

We shall prove that Sequence (2) is a normal realization of independent trials in which $p_i = \mu b_i$ (the b_i range over all possible m-

tuples consisting of the two symbols 0 and 1; $i = 1, 2, \ldots, 2^m$).
Let $\omega = (c_1 \ldots c_s)$ where $c_i = (\delta_0^{(i)} \ldots \delta_{m-1}^{(i)})$, and the $\delta_j^{(i)}$ are taken from an alphabet consisting of the symbols 0 and 1. We let $\tilde{N}_P(\omega)$ be the number of occurrences of ω prior to the Pth position of the caterpillar of Sequence (2). We let $\Omega = (\delta_1^{(1)} \ldots \delta_m^{(1)} \ldots \delta_1^{(s)} \ldots \delta_m^{(s)})$. We let $N_X(\Omega)$ be the number of occurrences of Ω prior to the Xth position of Sequence (1). It is clear that

$$\tilde{N}_P(\omega) < N_{Pm}(\Omega).$$

Hence

$$\overline{\lim_{P\to\infty}} \frac{\tilde{N}_P(\omega)}{P} < m\mu c_1 \ldots \mu c_s = m\mu\omega.$$

This means, according to the criterion, that Sequence (2) is a normal realization of independent trials. From this it follows, in particular, that if $\omega = c = (\delta_0 \ldots \delta_{m-1})$ where among the $\delta_0 \ldots \delta_{m-1}$ there are j ones and $m - j$ zeroes,

$$\lim_{P\to\infty} \frac{N_P(c)}{P} = p^j q^{m-j}. \tag{3}$$

Let $r_1 < r_2 < \ldots < r_k$ be nonnegative integers less than m. The quantity β_n constructed for Sequence (1) equals 1 if and only if there are ones at the r_1-th, ..., r_k-th positions of the element a_n [of Sequence (2)]. We let Σ be that set of m-tuples c consisting of symbols 0 and 1 for which there are ones at positions $r_1, \ldots, r_k$.

$$\mu\Sigma = p^k (p+q)^{m-k} = p^k.$$

In view of Equation (3) Sequence (2) possesses the property that the relative frequency of occurrence of Σ in it approaches p^k. But this is the definition of an admissible number.

Let there be given an admissible number

$$\varepsilon_0, \varepsilon_1, \ldots \tag{4}$$

We divide Sequence (4) into sections of length k

$$(\varepsilon_0 \varepsilon_1 \ldots \varepsilon_{k-1})(\varepsilon_k \ldots \varepsilon_{2k-1}) \ldots \tag{5}$$

Let Σ be some set of k-tuples characterized by the fact that at specific positions there are α ones for the elements Σ, while there are β zeroes at specific positions ($\alpha + \beta \leq k$). We let $T_l(\Sigma)$ indicate the number of times the set Σ is encountered prior to the lth position of Sequence (5). It is easy to establish (by induction with respect to β) that for an admissible number

$$\lim_{l\to\infty} \frac{T_l(\Sigma)}{l} = p^\alpha q^\beta. \tag{6}$$

Let Δ be a fixed k-termed combination consisting of symbols 0 and 1. Let $A \geq k$ be an increasing natural number. We take all possible A-tuples consisting of the symbols 0 and 1,

$$(b_1, \ldots, b_A).$$

Each such element is assigned a measure $p^j q^{A-j}$, where j is the number of ones among the terms $b_1 \ldots b_A$. To some set σ of different A-tuples we assign the measure $\mu\sigma$, equal to the sum of the measures of the elements entering into it.

Let us calculate how many times Δ is contained in the sequence

$$(b_1 \ldots b_k)(b_2 \ldots b_{k+1}) \ldots (b_{A-k+1} \ldots b_A). \tag{7}$$

We let σ_v be the set of those A-tuples for which Δ is contained in Sequence (7) v times.

Lemma (Markov). When $A \to \infty$

$$\sum_{v=1}^{\infty} v\mu\sigma_v = p^\alpha q^{k-\alpha} A + o(A),$$

where α is the number of ones in Δ.

Proof. (See [35], Chapter 6.)

Let $P = Al + r$, $0 \leq r \leq A - 1$,

$$N_P(\Delta) = \sum_{v=1}^{A-k+1} vT_l(\sigma_v) + O(lk) + O(r).$$

In view of Equations (6) and (8) we obtain

$$\overline{\lim_{P\to\infty}} \left| \frac{N_P(\Delta)}{P} - p^\alpha q^{k-\alpha} + o(1) \right| = O\left(\frac{k}{A}\right).$$

When A approaches ∞, we obtain

$$\lim_{P\to\infty}\frac{N_P(\Delta)}{P}=p^s q^{k-s},$$

which is what we were to establish.

§12. Completely Uniformly Distributed Sequences

Let s be a natural number. Consider the sequence of points

$$L_1, L_2, \dots, L_P, \dots, \tag{1}$$

lying in a unit cube of the s-dimensional space

$$L_P=(x_1^{(P)},\dots x_s^{(P)}).$$

Let Δ_s be any parallelepiped lying in a unit cube with edges parallel to the coordinate axes. We let $|\Delta_s|$ be its volume. We let $N_P(\Delta_s)$ be the number of points in Sequence (1) with numbers that do not exceed P, and which lie in Δ_s.

We say that the sequence (1) is uniformly distributed in a unit cube of an s-dimensional space if for any parallelepiped the relationship

$$\lim_{P\to\infty}\frac{N_P(\Delta_s)}{P}=|\Delta_s|. \tag{2}$$

holds.

If $s = 1$, we shall then say, as previously, that the sequence of numbers is uniformly distributed on the segment [01].

N.M. Korobov [1] has introduced the concept of a completely uniformly distributed sequcnce.

Consider an infinite sequence of numbers from the segment [01]

$$\alpha_1, \alpha_2, \alpha_3, \dots \tag{3}$$

We choose any natural number s and form the sequence of points in an s-dimensional unit cube

$$(\alpha_1, \alpha_2 \dots \alpha_s)(\alpha_2\alpha_3 \dots \alpha_{s+1})\dots \tag{4}$$

The Sequence (3) is called completely uniformly distributed if for any natural number s the sequence (4) is uniformly distributed

a unit cube of s-dimensional space.

More precisely speaking, N.M. Korobov has introduced an equivalent definition, e.g.:

The sequence of real numbers (3) is called completely uniformly distributed if for any natural number <u>s</u> and any set of integers $m_1, \ldots, m_s$, differing from 0, 0, ..., 0 the sequence

$$\beta_1, \beta_2, \beta_3, \ldots, \tag{5}$$

where $\beta_{\underline{l}} = \{m_1\alpha_{\underline{l}} + \ldots + m_s\alpha_{\underline{l}+s-1}\}$, $\underline{l} = 1, 2, 111$, is uniformly distributed on [01] ({} indicates a fraction).

The equivalents of these definitions follows from the following criteria for uniform distribution due to Weyl [2].

Lemma. <u>For the sequence of points</u>

$$(x_1^{(j)} \ldots x_s^{(j)}),\ j=1, 2, \ldots,\ 0 < x_i^{(j)} < 1,\ i=1, 2, \ldots, s$$

<u>to be uniformly distributed in a unit cube of an s-dimensional space, it is necessary and sufficient for the condition</u>

$$\sum_{j=1}^{P} e^{2\pi i (m_1 x_1^{(j)} + \cdots + m_s x_s^{(j)})} = o(P)$$

<u>to be satisfied for any set of integers</u> $m_1, m_2, \ldots, m_s$ <u>different from the set</u> 0, 0, ..., 0.

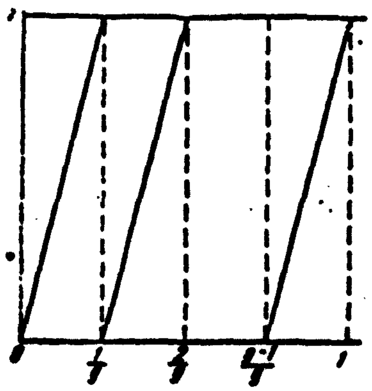

Fig. 1

In the next paragraph we shall construct a completely uniformly distributed sequence, and thus establish its existence.

By definition, any completely uniformly distributed sequence is uniformly distributed on the segment [01] (s = 1). There exist sequences, however, that are uniformly distributed on [01] and that are not completely uniformly distributed: for example, let $g \geq 2$ be a natural number; we select a real number α such

that the fractions $\{\alpha g^x\}$ are uniformly distributed (we have constructed such an α in this study. Consider the sequence of points in a unit square: $(\{\alpha g^x\}, \{\alpha g^{x+1}\})$, $x = 1, 2, \ldots$. In view of the self-evident relationship $\{\alpha g^{x+1}\} = \{\{\alpha g^x\} g\}$ we find that the points $(\{\alpha g^x\}, \{\alpha g^{x+1}\})$ are located within the unit square only on the lines drawn (Fig. 1), i.e., the sequence of points is not uniformly distributed within the unit square and what is more, the sequence $\{\alpha g^\alpha\}$, $x = 1, 2, \ldots$, is not completely uniformly distributed.

§13. Construction of a Completely Uniformly Distributed Sequence

There exist several methods for constructing completely uniformly distributed sequences.

The first example was proposed by N.M. Korobov [1]. In this example, the sequence is given as a sequence of fractions of some integral function whose argument runs through integral values. The integral function itself is given with the aid of a specially constructed power series. The proof that the sequence obtained is completely uniformly distributed makes use of an evaluation of trigonometric sums with polynomials, and is quite complex. In [25], N.M. Korobov gave another method for constructing a completely uniformly distributed sequence: the sequence is defined as a sequence of fractions $\{\alpha(x) q^x\}$ where $x = 1, 2, \ldots$, and $\underline{q}$ is an integer ≥ 2, $\alpha(x)$ is a specially constructed integral function. A simpler technique than that used in [1] is used to prove that the sequence constructed is completely uniformly distributed; this construction, however, cannot be considered to be simple.

The construction presented below was carried out by L.P. Starchenko [27].

This construction is technically simple, but it is based on a profound property of transcendental numbers. L.P. Starchenko [37] ha

given still another construction for a completely uniformly distributed sequence. We shall discuss this method in §15 of this monograph.

Let p_r — r-e be a prime number. We let $n_k = [e^{(\ln k)^3}] + 1$, $k = 1, 2, \ldots$. We consider the sequence of real numbers

$$\begin{gathered}\{\ln 2\},\ \{2\ln 2\},\ \ldots,\ \{n_1 \ln 2\},\ \{\ln 2\},\ \{\ln 3\},\ \ldots,\ \{n_2 \ln 2\},\\ \{n_2 \ln 3\},\ \{\ln 2\},\ \{\ln 3\},\ \ldots,\ \{\ln p_r\},\ \ldots,\ \{n_r \ln 2\},\ \ldots,\ \{n_r \ln p_r\}\ \ldots\end{gathered} \tag{1}$$

We designate this sequence as

$$\alpha_1,\ \alpha_2,\ \alpha_3,\ \ldots.$$

Theorem. Sequence (1) is completely uniformly distributed.

Proof. Let s be a fixed natural number, $m_1, \ldots, m_s$ be integers not all equal to zero and $m = \max(m_1, \ldots, m_s)$. Consider the trigonometric sum

$$S = \sum_{l=1}^{P} e^{2\pi i (m_1\alpha_l + \ldots + m_s\alpha_{l+s-1})}.$$

We let $T_j = n_1 + 2n_2 + \ldots + jn_j$. Let

$$T_k + \lambda(k+1) \leqslant P < T_k + (\lambda+1)(k+1).$$

We write the sum in the form

$$\begin{aligned} |S| < {} & \Bigg|\sum_{l=1}^{T_s} e^{2\pi i (m_1\alpha_l + \cdots + m_s\alpha_{l+s-1})} + \\ & + \sum_{j=s+1}^{k}\Bigg[\Bigg\{\sum_{l=1}^{n_j}\Bigg(\sum_{r=1}^{j-s+1} e^{2\pi i l (m_1 \ln p_r + \cdots + m_s \ln p_{r+s-1})}\Bigg) + \\ & + \sum_{r=j-s+2}^{j} e^{2\pi i\, (l(m_1 \ln p_r + \cdots + m_{j-r+1}\ln p_j) + (l+1)(m_{j-r+2}\ln 2 + \cdots + m_s \ln p_{r+s-j-1}))} + \\ & + O(1)\Bigg\}\Bigg]\Bigg| + \Bigg|\sum_{l=1}^{\lambda}\Bigg[\sum_{r=1}^{k+1-s+1} e^{2\pi i l (m_1 \ln p_r + \cdots + m_s \ln p_{r+s-1})} + \\ & + \sum_{r=k+3-s}^{k+1} e^{2\pi i\, (l(m_1 \ln p_r + \cdots + m_{k+1-r+1}\ln p_{k+1}) + (l+1)(m_{k+2-r}\ln 2 +} \\ & \qquad {}^{+ m_s \ln p_{r+s-(k+1)-1}))} + O(k)\Bigg]\Bigg|. \end{aligned}$$

Since s is fixed, the first sum is O(1). We now change the order of summation over the indices l and r. It is possible to make use of the well-known condition

$$\left|\sum_{x=P_1}^{P_2} e^{2\pi i \alpha x}\right| < \frac{1}{2(\alpha)},$$

where (α) is the distance of α to the nearest integer

$$|S| < \sum_{j=s+r}^{k+1}\left(\sum_{r=1}^{j-s+1}\frac{1}{(m_1 \ln p_r + \cdots + m_s \ln p_{r+s-1}} + \right.$$

$$+ \sum_{r=s+j-s}^{k+1}\frac{1}{([m_1 \ln p_r + \cdots + m_{j-r+1} \ln p_j] / + (j+1)(m_{j-r+s} \ln r + \cdots + m_s \ln p_{r+s-j-1}))} +$$

$$\left. + O(k)\right).$$

We shall in addition require the following lemma, which gives a quantitative result of the well-known fact that numbers of the form e^a are irrational, where a is an integer [26].

Lemma 1. Let f, f_1, and a be integers, $f > f_1 \geq 1$, $a < 2 \ln f$, $m = [10 \ln f]$, $n = [3 \ln (m + 1)] + 1$. Then

$$\left|\ln \frac{f}{f_1} - a\right| > \frac{1}{2^{m} f^{m-1} e^{3(m+1)n}}.$$

An obvious consequence of this lemma is Lemma 2:

Lemma 2. Let $f \neq f_1$, f and f_1 be natural numbers. Let $H = \max (f, f_1)$. There exists a constant $c > 0$ such that

$$\left(\ln \frac{f}{f_1}\right) > \frac{1}{e^{c \ln^2 H}}.$$

Let

$$m_1 \ln p_{l_1} + \cdots + m_s \ln p_{l_s} = \ln \frac{f}{f_1} \quad (l_s < j).$$

Then

$$H < p_j^{m_s}.$$

It is known that $p_j \leq j \ln j$ and thus $H \leq j^{2ms}$. Hence (c_1, c_2, c_3, c_4 are positive constants)

$$(S) < \sum_{j=1}^{k+1} j^{c_1 (ms)^3 \ln^2 j} + O(k) < e^{c_2 (ms)^3 \ln^3 (k+1)} < e^{c_3 (ms)^3 \ln^3 k}.$$

But $P \geq n_k > e^{(\ln k)^3}$. Therefore

$$|S| < e^{c_4 (ms)^3 (\ln P)^{\frac{1}{3}}} = o(P).$$

By the Weyl criterion, Sequence (1) is completely uniformly distributed.

§14. N.M. Korobov's Theorem

N.M. Korobov (see [1], page 217) has proved the following theorem.

Theorem. If a sequence of real numbers from the segment [01]

$$\alpha_1, \alpha_2, \ldots, \alpha_p \qquad (1)$$

is completely uniformly distributed, then a sequence of the first base-g terms

$$[\alpha_1 g], \ [\alpha_2 g], \ldots, \ [\alpha_p g], \ldots \qquad (2)$$

is a normal sequence of terms 0, 1, ..., g − 1.

Proof. In fact, for any natural number s, the presence of the combination $\Delta_s = (\delta_1 \ldots \delta_s)$ in the caterpillar for Sequence (2) is equivalent to the fact that the elements of the corresponding term in the caterpillar for Sequence (1) will fall within the interval $\left[\frac{\delta_i}{g} \frac{\delta_{i+1}}{g}\right)$, i = 1, 2, ..., s. Since Sequence (1) is completely uniformly distributed,

$$\lim_{p\to\infty} \frac{N_p(\Delta_s)}{p} = \frac{1}{g^s},$$

i.e., Sequence (2) is normal.

Remark. Since there exist methods for constructing completely uniformly distributed sequences, N.M. Korobov's theorem yields methods for constructing normal sequences of symbols.

§15. The Converse of N.M. Korobov's Theorem

L.P. Starchenko has proved the statement that is the converse of N.M. Korobov's theorem [37, 38].

Let $g \geq 2$ be a natural number, and

$$a_1, a_2, a_3, \ldots \qquad (1)$$

be a given normal sequence composed of the symbols 0; 1, ..., g − 1.

Using the method given in §8, we construct an infinite number of sequences

$$a_1^{(i)}, a_2^{(i)}, a_3^{(i)}, \ldots$$

($i = 1, 2, 3, \ldots$) such that for any natural number l the system of sequences

$$\begin{array}{l} \varepsilon_1, \quad \varepsilon_2, \quad \varepsilon_3, \ldots \\ \varepsilon_1^{(1)}, \quad \varepsilon_2^{(1)}, \quad \varepsilon_3^{(1)}, \ldots \\ \cdots\cdots\cdots\cdots \\ \varepsilon_1^{(l-1)}, \varepsilon_2^{(l-1)}, \varepsilon_3^{(l-1)}, \ldots \end{array} \tag{2}$$

is mutually normal.

Theorem. The sequence of real numbers

$$\alpha_1, \alpha_2, \alpha_3, \ldots, \tag{3}$$

where $\alpha_j = 0, \varepsilon_j\varepsilon_j^{(1)}\varepsilon_j^{(2)}\ldots$ [i.e., $\alpha_j = (\varepsilon_j/g) + (\varepsilon_j^{(1)}/g^2) + \ldots$] is a completely uniformly distributed sequence.

Proof. By the definition of completely uniformly distributed sequence, it is necessary to show that for any natural number s, the sequence of points

$$Q_1, Q_2, Q_3, \ldots, \tag{4}$$

where

$$Q_k = (\alpha_k, \alpha_{k+1}, \ldots, \alpha_{k+s-1}),$$

is uniformly distributed in an s-dimensional unit cube.

We take an arbitrary natural number l and consider the sequence of points

$$Q_1^{(l)}, Q_2^{(l)}, Q_3^{(l)}, \ldots, \tag{5}$$

where $Q_k^{(l)} = (\alpha_k^{(l)}, \alpha_{k+1}^{(l)}, \ldots, \alpha_{k+s-1}^{(l)})$, while the number $\alpha_j^{(l)}$ is defined as follows:

$$\alpha_j^{(l)} = 0, \varepsilon_j\varepsilon_j^{(1)}\ldots\varepsilon_j^{(l-1)}.$$

Since the system of sequences (2) is mutually normal, every possible point $Q_k^{(l)}$ in Sequence (5) will be encountered with an asymptotic frequency of $1/g^{ls}$.

But when we find a point $Q_k^{(l)}$, it means that the corresponding point Q_k falls within a cube defined by the system of inequalities

$$0, \varepsilon_j\varepsilon_j^{(1)}\ldots\varepsilon_j^{(l-1)} < \alpha_{k+x} < 0, \varepsilon_j\varepsilon_j^{(1)}\ldots\varepsilon_j^{(l-1)} + \frac{1}{g^l}, \quad x = 0, 1, \ldots, s-1. \tag{6}$$

Thus, the number of times that points of Sequence (4) fall within any cube of the form (6) will have an asymptotic frequence of $1/g^{\underline{l}s}$, i.e., it will equal the volume of cube (6). Since any parallelepiped lying within a unit cube of s-dimensional space may with any degree of accuracy be approximated by cubes of the form (6) (taking large enough $\underline{l}$) the theorem is proved.

§16. A Sequence Completely Distributed Over the Function F(x)

Let there be given a distribution function $F(x)$. We associate the distribution function $F(x)$ and the measure μ on a line that is also in an s-dimensional arithmetic space, as follows: the measure of the segment $\Delta = [x_1 x_2]$, where x_1 and x_2 are points of continuity of the function $F(x)$ equals $\mu\Delta = F(x_2) - F(x_1)$; if Δ_s is a parallelepiped in s-dimensional space whose projections on the coordinate axes are the segments $\Delta^{(1)}, \ldots, \Delta^{(s)}$ with ends that are points of continuity of the function $F(x)$, then

$$\mu\Delta_s = \mu\Delta^{(1)} \ldots \mu\Delta^{(s)}.$$

Consider the sequence of real numbers

$$\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_p, \ldots \tag{1}$$

We take a natural number $\underline{s}$ and form the sequence of tuples

$$(\alpha_1 \alpha_2 \ldots \alpha_s)(\alpha_2 \alpha_3 \ldots \alpha_{s+1}) \ldots (\alpha_p \ldots \alpha_{p+s-1}) \ldots \tag{2}$$

We write Sequence (2) as a sequence of points in an s-dimensional arithmetic space

$$Q_1, Q_2, Q_3, \ldots, Q_p, \ldots, \tag{2'}$$

where $Q_i = (\alpha_i, \alpha_{i+1}, \ldots, \alpha_{i+s-1})$, $i = 1, 2 \ldots$.

We shall henceforth let Δ_s stand for the parallelepiped $\Delta_s = (\Delta^{(1)} \ldots \Delta^{(s)})$ in an s-dimensional arithmetic space such that the ends of the interval $\Delta^{(1)}, \ldots, \Delta^{(s)}$ are points of continuity for the function $F(x)$.

We let $N_P(\Delta_s)$ be the number of points in Sequence (2) prior to

the Pth number that lie within Δ_s.

We call Sequence (1) completely uniformly distributed over the function F(x) if for any natural number s and any parallelepiped Δ_s

$$\lim_{P\to\infty}\frac{N_P(\Delta_s)}{P}=\mu\Delta_s.$$

Special cases:

1) Let

$$F(x)=\begin{cases}0 & x<0,\\ \frac{1}{g} & 0<x<\frac{1}{g},\\ \dots\dots & \\ 1 & \frac{g-1}{g}<x.\end{cases}$$

If Sequence (1) consists of the numbers 0, 1, ..., g − 1, this is a normal sequence of terms.

2) Let

$$F(x)=\begin{cases}0 & x<0,\\ x & 0<x<1,\\ 1 & 1<x.\end{cases}$$

If Sequence (1) is made ~~up of real~~ numbers taken from the segment [01], then the sequence is completely uniformly distributed on the segment [01].

3) Let there be given two positive numbers p and q such that p + q = 1 and

$$F(x)=\begin{cases}0 & x<0,\\ q & 0<x<1,\\ 1 & 1<x.\end{cases} \qquad (3)$$

Let Sequence (1) be composed of the numbers 0 and 1. The sequence completely distributed over the function F(x) defined by Equations (3) is a Bernoulli-normal sequence of symbols.

4) Let $a > 0$, $p(m) = (a^m/m!)e^{-a}$, where $m \geq 0$ is an integer.

Let us consider the following distribution function

$$F(x)=\begin{cases}0 & x<0,\\ \sum_{m\leq x} p(m) & x>0.\end{cases} \qquad (4)$$

If Sequence (1) is made up of negative integers and is completely

tributed over the function $F(x)$ defined by Eqs. (4), we call such a sequence completely Poisson-law distributed.

5) Let $F(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{z^2}{2}}dz$. If Sequence (1) is completely distributed over this function $F(x)$ then Sequence (1) is said to be completely Gauss-law (or normally) distributed.

§17. Construction of a Sequence Completely Distributed With Respect to a Function F(x)

If there is a sequence

$$\alpha_1, \alpha_2, \ldots, \alpha_p, \ldots, \tag{1}$$

completely uniformly distributed on the segment [01], is is then possible to construct a sequence completely distributed with respect to a function $F(x)$. This method generalizes the theorem of N.M. Korobov (§14 of this study).

Let us construct a sequence of real numbers

$$\beta_1, \beta_2, \ldots, \beta_p, \ldots, \tag{2}$$

where β_j is defined by the equation

$$F(\beta_j) < \alpha_j < F(\beta_j + 0).$$

We shall prove that Sequence (2) is completely distributed with respect to function $F(x)$.

We take $s \geq 1$ and form a sequence of points in a unit cube of s-dimensional space

$$Q_1, Q_2, Q_3, \ldots, Q_p, \ldots, \tag{3}$$

where $Q_p = (\alpha_p \ldots \alpha_{p+s-1})$, and the sequence of points in s-dimensional space

$$L_1, L_2, \ldots, L_p, \ldots,$$

where

$$L_p = (\beta_p \ldots \beta_{p+s-1}).$$

We take the parallelepiped $\Delta_s = (\Delta^{(1)}, \ldots, \Delta^{(s)})$ where $\Delta^{(i)} = (a^{(i)} b^{(i)})$, $i = 1, 2, \ldots, s$; $a^{(1)}, b^{(1)}, \ldots, a^{(s)}, b^{(s)}$ are points

of continuity of the function $F(x)$. The number of points of Sequence (4) prior to the Pth number that lie within parallelepiped Δ_s equals the number of points of Sequence (3) prior to the Pth number that lie within the parallelepiped $\bar{\Delta} = (\bar{\Delta}^{(1)}, \ldots, \bar{\Delta}^{(s)})$, where $\bar{\Delta}^{(i)} = (c^{(i)}, d^{(i)})$, $i = 1, 2, \ldots, s$; $c^{(i)} = Fa^{(i)})$, $d^{(i)} = F(b^{(i)})$. But since Sequence (1) is completely uniformly distributed within the unit cube, when $P \to \infty$ this number is equivalent to

$$P|\Delta| = P\prod_{i=1}^{s}(F(b^{(i)}) - F(a^{(i)})) = P\mu\Delta_s.$$

Thus, for Sequence (2)

$$\lim_{P\to\infty}\frac{N_P(\Delta_s)}{P} = \mu\Delta_s,$$

which was what we set out to prove.

§18. Measure in a Space of Infinite Sequences of Symbols

We have made use of I.I. Pyatetskiy's criteria to prove that a constructed sequence is a normal sequence of symbols. For constructions to come, we shall also need theorems that extend the criteria of I.I. Pyatetskiy. There are difficulties involved in proving such theorems by the method presented in §4 (the calculations become cumbersome). There is a more flexible method of proof, however, based on the study of the metric properties of dynamic systems. In the following sections, we shall give such a proof for a Pyatetskiy criterion.

We will require some information from measure theory [39]. In the classical theory of functions, a space with a measure is a finite-dimensional Euclidean space. We shall use a different example of a space with a measure.

Let $g \geq 2$ be a natural number.

Consider the set of all infinite sequences made up of the symbols $0, 1, \ldots, g-1$

$\delta_1\delta_2 \ldots \delta_n \ldots$

We call this set the space R, and we let the letter p designate the elements of this set.

We shall consider certain sets of subsets of R (following Halmos [39] we shall call such sets classes of sets).

Let s be any natural number. The set M of sequences in which the first s symbols $\delta_1, \ldots, \delta_s$ are fixed shall be called elementary cylinder sets, and we shall designate them as $M(\delta_1 \ldots \delta_s)$. The space R and the empty set 0 are also elementary cylinder sets.

We note that two elementary cylinder sets are either disjoint or equal, which will simplify our study.

We call a set that may be written as a finite set-theoretical sum of elementary cylinder sets a cylinder set.

Lemma 1. Cylinder sets form an algebra.*

Proof. We shall first show that the intersection of two cylinder sets M and M' is also a cylinder set. Let

$$M = \bigcup_i M_i, \quad M' = \bigcup_j M'_j,$$

where M_i and M'_j are elementary cylinder sets

$$M \cap M' = \bigcup_i \bigcup_j M_i \cap M'_j.$$

But the $M_i \cap M'_j$ are elementary cylinder sets.

We shall now show that the complement of a cylinder set is a cylinder set. We let CM be the complement of the set M. The complement of the set $M(\delta_1 \ldots \delta_s)$ is a sum of sets of the form $M(\tau_1 \ldots \tau_s)$, in which $\tau_1 \ldots \tau_s$ form all possible combinations of symbols with the exception of $\delta_1 \ldots \delta_s$, i.e., the complement is the sum of a finite number of elementary cylinder sets. The formula

$$C(M_1 \cup M_2) = CM_1 \cap CM_2.$$

is self-evident. From this formula and from the fact that the inter-

section of two cylinder sets is a cylinder set it follows that the complement of a cylinder set is a cylinder set.

Using the rule

$$1.\ \mu(R)=1;\quad 2.\ \mu(0)=0;\quad 3.\ \mu(M(\delta_1\ldots\delta_s))=\frac{1}{g^s}.$$

we introduce a function whose argument is an elementary cylinder set M, $\mu(M)$. Let us study the properties of this function:

Lemma 2. Let there be a cylinder set M

$$M=\bigcup M_i,$$

where the M_i are elementary cylinder sets.

There exists a system of disjoint elementary cylinder sets γ_1, γ_2, ..., γ_N such that

$$1.\ M=\bigcup\gamma_i.$$
$$2.\ \text{If}\quad M_i\cap\gamma_j\neq 0\ \text{then}\ \gamma_j\subseteq M_i.$$

3. μM_i equals the sum $\mu\gamma_j$ of those γ_j that have points in common with M_i.

Proof. Let $M_i=(\delta_1^{(i)}, \ldots, \delta_{s_i}^{(i)})$ and $s=\max s_i$. If $s_i < s$, we represent M_i as a sum g^{s-s_i} of elementary cylinder sets in which the first s terms coincide with $\delta_1^{(i)}, \ldots, \delta_{s_i}^{(i)}$, while the remaining $s-$ $-s_i$ are varied in all possible ways. We include all of these sets in the system γ_1, γ_2, ... If $s=s_i$, then we place the M_i in the system γ_1, γ_2, ... One representative from each group of identical sets is left in the system of sets thus obtained. We obtain the system of sets required. In fact, the sets γ_i are disjoint. Let us check all the assertions of the lemma in turn:

1. By construction, all the $M_i=\bigcup\gamma_{j(i)}$, while $M=\bigcup M_i$.

2. Two cylinder sets are either disjoint, equal, or one is contained in the other (here the set contained is that with the smaller number of fixed symbols).

3. Those γ_1, ..., γ_N that enter into the representation of M_i

a sum g^{s-s_i} of elementary cylinder sets intersect with M_i. The equation $g^{s-s_i}(1/g^s) = 1/g^{s_i}$ shows that μM_i equals the sum $\mu\gamma_j$ of those γ_j that have points in common with M_i.

Lemma 3. Let $\{M_i\}$ and $\{F_j\}$ be two finite systems of elementary cylinder sets; then

$$\bigcup M_i \subset \bigcup F_j$$

and the M_i are mutually disjoint. Then

$$\sum \mu M_i \leq \sum \mu F_j.$$

Proof. Consider the sum

$$M = \bigcup M_i \bigcup F_j.$$

We apply the preceding lemma to the set M. Since the sets M_i are disjoint, the same γ_k cannot be contained in two different M_i. Consequently those γ_k contained in $\bigcup M_i$ can be separated into groups with group T_i containing those γ_k contained in M_i. The groups T_i are disjoint

$$\sum \mu M_i = \sum_i \Big(\sum_{T_i} \mu\gamma_k\Big) = \sum_T \mu\gamma_k,$$

where T is the set of all γ_k contained in $\bigcup M_i$. We now distribute the sets γ_k contained in T among other groups. In particular, let S_1 be the set of all γ_k (contained in T) that are contained in F_1, let S_2 be the set of all γ_k contained in F_2 but not in F_1, and let S_3 be the set of γ_k contained in F_3 but not in $S_1 \cup S_2$... . Then:

$$\sum_{S_j} \mu\gamma_k \leq \mu F_j.$$

Since μF_j (by Lemma 2) equals the sum $\mu\gamma_k$ of all γ_k included in F_j. This means that

$$\sum \mu M_i = \sum_T \mu\gamma_k = \sum_j \Big(\sum_{S_j} \mu\gamma_k\Big) \leq \sum \mu F_j.$$

The lemma is proved.

Corollary. If some set M can be written as the finite sum of dis-

joint elementary cylinder sets in two ways

$$M = \bigcup M_i, \quad M = \bigcup F_j,$$

then

$$\sum \mu M_i = \sum \mu F_j.$$

In fact, $\bigcup M_i \subset \bigcup F_j$ and $\bigcup F_j \subset \bigcup M_i$. Thus by Lemma 3

$$\sum \mu M_i = \sum \mu F_j.$$

We now extend the function $\mu(M)$ from elementary cylinder sets to cylinder sets.

Definition: Let M be a cylinder set

$$M = \bigcup_i M_i,$$

where $\bigcup_i M_i$ is the sum of a finite number of disjoint elementary cylinder sets. We say that $\mu M = \sum_i \mu M_i$.

In view of the corollary that has been proven, this definition is consistent.

Lemma 4. Let $F = \bigcup_{i=1}^{s} M_i$, where $M_1, \ldots, M_s$ are disjoint cylinder sets. Then

$$\mu F = \sum_{i=1}^{s} \mu M^i.$$

Proof. We decompose each M_i into disjoint elementary cylinder sets

$$M_i = \bigcup M_{ij}.$$

Then

$$\mu M_i = \sum \mu M_{ij}.$$

Since the M_i are disjoint, the M_{ij} are also disjoint.

$$F = \bigcup_i \bigcup_j M_{ij},$$

$$\mu F = \sum_i \sum_j M_{ij}.$$

Thus

$$\mu F = \sum \mu M_i.$$

We shall now prove that the function μ, in a rather strange form, is countably additive.

Lemma 5. The sum of a countable number of nonempty disjoint cylinder sets will not be a cylinder set.

Proof. Let the cylinder set F be the sum of an infinite number of nonempty disjoint cylinder sets $F = \cup F_i$. We decompose each F_i into disjoint elementary cylinder sets. By Lemma 1, CF is a cylinder set; we represent CF as the sum of disjoint elementary cylinder sets. We obtain a representation for R as the sum of a countable number of elementary cylinder sets

$$R = \bigcup_{i=1}^{\infty} M_i.$$

Let the $M_i^{(a_1)}$ be those terms in which the first element is a_1. In general, we will denote by $M_i^{(a_1 \ldots a_k)}$ those terms of M_i in which the first k elements are respectively $a_1, a_2, \ldots, a_k$. Clearly,

$$M(j) = \bigcup_i M_i^{(j)}, \quad j = 0, 1, \ldots, g-1.$$

There is at least one j, $j = j_1$ such that the sum $\cup M_i^{(j_1)}$ contains an infinite number of terms. In addition, it can be seen that there is at least one j_2 such that the right side of equation

$$M(j_1, j_2) = \cup M_i^{(j_1, j_2)}$$

contains an infinite number of terms. We continue this argument indefinitely. Let α be the sequence

$$\alpha = j_1, j_2, j_3, \ldots$$

Since $\alpha \in R$, α is contained in one of the M_i. Let α be contained in $M(j_1 j_2 \ldots j_s)$. Clearly $M(j_1 j_2 \ldots j_s)$ is the only term in the sum $\cup M_i$ in which the first s symbols are $j_1, j_2, \ldots, j_s$. Thus in the sum

$$M(j_1, j_2, \ldots, j_s) = \cup M_i^{(j_1 j_2 \ldots j_s)}$$

there cannot be an infinite number of terms in the right side. This

contradiction proves the lemma.

The function μ is nonnegative, defined on an algebra, countably additive, and $\mu(0) = 0$, i.e., the function μ is by definition a measure ([39], page 34).

As we know (for example, from [39], page 59) there exists a uniquely defined minimal σ algebra $\mathfrak{B}$ containing the algebra of cylinder sets to which the measure μ is uniquely extended.

We will call the sets of the σ algebra $\mathfrak{B}$ measurable. We call the function $f(p)$ measurable if it takes on real values and if for any real $\underline{c}$ the set of points $\underline{p}$ in which $f(p) \leq c$ is measurable.

The theory of the Lebesgue integral is similar to the theory of functions of a real variable (see [40]).

<u>§19. A Dynamic System in a Space of Sequences of Terms</u>

We define the family of transformations T^k, $k = 0, 1, \ldots$ in R. Let

$$p = \varepsilon_1, \varepsilon_2, \varepsilon_3, \ldots,$$

$$T^k p = \varepsilon_{k+1}, \varepsilon_{k+2}, \varepsilon_{k+3}, \ldots$$

The transformation T is not one-to-one. The complete preimage T^{-1} of a point $\underline{p}$ consists of the $\underline{g}$ points $k\varepsilon_1\varepsilon_2 \ldots \varepsilon_k$; $k = 0, 1, \ldots, g - 1$.

Lemma 1. <u>The measure</u> μ <u>is an invariant measure, i.e., the complete preimage of a measurable set is measurable, and</u> $\mu T^{-1} M = \mu M$.

Proof. It is sufficient to show that for an elementary cylinder set M $\mu T^{-1} M = \mu M$. The complete preimage of the elementary cylinder set $M(\delta_1 \ldots \delta_s)$ consists of $\underline{g}$ disjoint elementary cylinder sets

$$M_i(\varepsilon_i \delta_1 \ldots \delta_s), \quad \varepsilon_i = 0, 1, \ldots, g - 1,$$

$$\mu M_i = \frac{\mu M}{g}, \quad \sum \mu M_i = g \frac{\mu M}{g} = \mu M.$$

The space R, together with the invariant measure μ and the family of transformations T^k, forms what is called a dynamic system.

A set M is called invariant if $T^{-1} M = M$. Sets of all sequences

that are periodic following a certain term are examples of invariant sets.

Lemma 2. <u>The space</u> R <u>cannot be represented as the sum of two measurable invariant disjoint sets with positive measure</u>.

Proof. We assume that we are able to represent R as the sum of two sets without common points of positive measure, $R = U_1 \cup U_2$. We let $\eta = \mu U_1$, $0 < \eta < 1$. Let $\chi(p)$ be the characteristic function of the set U_1. Let any combination $\varepsilon_1 \varepsilon_2 \ldots \varepsilon_n$ of n terms be given and let

$$p = \delta_1 \delta_2 \ldots$$

be some point in R. Since the sequence

$$\varepsilon_1 \varepsilon_2 \ldots \varepsilon_n \delta_1 \delta_2 \ldots$$

is one of the nth-order preimages of point p and U_1 is an invariant set,

$$\chi(\varepsilon_1 \varepsilon_2 \ldots \varepsilon_n p) = \chi(p).$$

The measure of the intersection of U_1 with the elementary cylinder set $M(\varepsilon_1 \ldots \varepsilon_n)$ is $\int_{M(\varepsilon_1 \ldots \varepsilon_n)} \chi(p)\, d\mu$. Clearly

$$\int_{M(\varepsilon_1 \ldots \varepsilon_n)} \chi(p)\, d\mu = \frac{1}{g^n} \int_R \chi(p)\, d\mu = \frac{1}{g^n} \mu U_1 = \frac{1}{g^n} \eta.$$

Assume $\varepsilon > 0$ and $1 - \eta > \varepsilon$ (both inequalities are strict). By a theorem similar to the theorem on accumulation points (see [40], page 286), the set U_1, as a set with positive measure, should have an accumulation point ϑ_0.

$$\vartheta_0 = \tau_1 \tau_2 \ldots,$$

i.e., for every $\varepsilon > 0$ it is possible to find a δ_0 such that whatever elementary cylinder set Δ with $\mu\Delta < \delta_0$ we take that contains the point ϑ_0

$$\frac{\mu(U_1 \cap \Delta)}{\mu\Delta} > 1 - \varepsilon.$$

We take an n so large that $1/g^n < \delta$, while for Δ we take the elementary

cylinder set $M(\tau_1\tau_2\ldots\tau_n)$. On the one hand

$$\mu(U_1\cap\Delta)=\frac{1}{g^n}\eta,$$

on the other hand

$$\mu(U_1\cap\Delta)>(1-\varepsilon)\frac{1}{g^n}.$$

This yields $\eta > 1-\varepsilon$, which contradicts the condition $1-\eta>\varepsilon$. This proves the lemma.

§20. Birkhoff-Khinchin Theorem

Let R be a set of points and p the points in R. Let $\mu(\mu R = 1)$ be a normalized measure defined on a σ algebra of sets $\mathfrak{B}$ in R.

Let the family of transformations T^k, $k = 0, 1, 2, \ldots$, R be defined on itself such that for any integers $k_1 \geq 0$ and $k_2 \geq 0$

$$T^{k_1+k_2}p = T^{k_1}(T^{k_2}p),\ p\in R.$$

Generally speaking, we do not assume that the generating transformation T is one-to-one. We let $T^{-1}A$ be the complete preimage of some set A.

We shall assume that the complete preimage of every set of lies in $\mathfrak{B}$ and that

$$\mu T^{-1}A=\mu A$$

(measure in variants).

We shall call the space R together with the measure μ and the family of transformations T^k, $k = 0, 1, \ldots$ a dynamic system.

The set E is called invariant if

$$T^{-1}E=E.$$

If R cannot be represented as the sum of two disjoint sets (in $\mathfrak{B}$), both of positive measure, we then say that the dynamic system is indecomposable (or ergodic).

We shall need to use the so-called Birkhoff-Khinchin ergodic theorem. This theorem was proved by Birkhoff [41] and Khinchin [42]

for the case of one-to-one transformations. F. Riez* [43] has extended this theorem to transformations that are not one-to-one. The proof given in [43] is concise and I shall therefore give it here in its entirety.

Theorem (part 1). Let μ be an invariant measure. Let there be a function $\varphi(p)$ that is absolutely summable over the measure μ (i.e., the integral $\int_R |\varphi(p)| d\mu$ exists). Then the limit

$$\lim_{n\to\infty} \frac{\varphi(p)+\varphi(Tp)+\ldots+\varphi(T^{n-1}p)}{n} = \psi(p).$$

$$\int_R \psi(p)\, d\mu = \int_R \varphi(p)\, d\mu.$$

exists for nearly all points $p \in R$ (with respect to the measure μ).

Proof. We let

$$\underline{\lim}_{n\to\infty} \frac{\varphi(p)+\varphi(Tp)+\ldots+\varphi(T^{n-1}p)}{n} = \psi_*(p).$$

$$\overline{\lim}_{n\to\infty} \frac{\varphi(p)+\varphi(Tp)+\ldots+\varphi(T^{n-1}p)}{n} = \psi^*(p).$$

Consider a set of intervals on the real line with rational end points $\{\alpha_n \beta_n\}$ (this is a denumerable set). Consider the set V_n of points p such that for these points

$$\psi_*(p) < \alpha_n < \beta_n < \psi^*(p).$$

If $\mu V_n = 0$ ($n = 1, 2 \ldots$), then letting $V' = \bigcup_{n=1}^{\infty} V_n$ we find that $\mu V' = 0$. If $p \in R \setminus V'$ it is impossible to insert any interval with rational end points such as $\psi_*(p) = \psi^*(p)$ between $\psi_*(p)$ and $\psi^*(p)$. Thus to prove the theorem, we must assume the existence of the set V_n (i.e., of the two rational numbers α_n and β_n) such that $\mu V_n = 0$, and this must lead to a contradiction. Let us make a change in notation: we shall write S in place of V_n, and α and β in place of α_n and β_n. Thus S is the set with $\mu S > 0$ for the points of which

$$\underline{\lim}_{n\to\infty} \frac{\varphi(p)+\ldots+\varphi(T^{n-1}p)}{n} < \alpha < \beta < \overline{\lim}_{n\to\infty} \frac{\varphi(p)+\ldots+\varphi(T^{n-1}p)}{n}.$$

We note that the set S is an invariant set. In fact, let $p' \in S^{-1}$;

then $p = Tp' \in S$ and

$$\frac{\varphi(p')+\varphi(Tp')+\dots+\varphi(T^{n-1}p')}{n} = \frac{\varphi(p')}{n} + \frac{n-1}{n}\,\frac{\varphi(p)+\dots+\varphi(T^{n-2}p)}{n-1},$$

$$\varliminf_{n\to\infty}\frac{\varphi(p')+\varphi(Tp')+\dots+\varphi(T^{n-1}p')}{n} = \varliminf_{n\to\infty}\frac{\varphi(p)+\dots+\varphi(T^{n-2}p)}{n-1} < \alpha,$$

$$\varlimsup_{n\to\infty}\frac{\varphi(p')+\varphi(Tp')+\dots+\varphi(T^{n-1}p')}{n} = \varlimsup_{n\to\infty}\frac{\varphi(p)+\dots+\varphi(T^{n-2}p)}{n-1} > \beta.$$

We shall later require the following lemma:

Lemma. *Let* n *real numbers* $a_1, a_2, \dots, a_n$ *and an integer* $m < n$ *be given. Let us consider all sums formed by sequences of numbers having a number of terms less than or equal to* m

$$a_k + a_{k+1} + \dots + a_l,$$

that are greater than zero. We shall say that the numbers in our sequence a_k *are selected if they figure in at least one of these sums as the first term. The sum of the selected numbers will be greater than zero.*

Proof. Let a_{k_1} be the smallest selected number. Let $a_{k_1} + \dots + a_{l_1}$ be the shortest positive sum beginning with a_{k_1}. Then all of its terms will be selected numbers, e.g.: $a_{\tau_1} + \dots + a_{l_1} > 0$. In fact, if $a_{\tau_1} + \dots + a_{l_1} \le 0$, then since $a_{k_1} + \dots + a_{l_1} > 0$, it follows that $a_{k_1} + \dots + a_{\tau_1 - 1} > 0$, which contradicts the assumption that $a_{k_1} + \dots + a_{l_1}$ is the shortest positive sum beginning with a_{k_1}. Let a_{k_2} be the smallest selected number larger than l_1 and let $a_{k_2} + \dots + a_{l_2}$ be the shortest positive sum beginning with a_{k_2}. It may be seen that all of the terms of this sum are selected numbers. In this manner, we can find all the selected numbers. Their sum will equal

$$\sum_i (a_{k_i} + \dots + a_{l_i}).$$

This sum is positive since every $a_{k_i} + \dots + a_{l_i} > 0$. The lemma is proved.

If $p \in S$ there exists an n such that

$$\frac{\varphi(p)+\varphi(Tp)+\ldots+\varphi(T^{n-1}p)}{n}>\beta$$

or

$$\varphi(p)-\beta+\varphi(Tp)-\beta+\ldots+\varphi(T^{n-1}p)-\beta>0.$$

We let $\underline{l}$ stand for the smallest of these values of $\underline{n}$; $\underline{l} = \underline{l}(p)$. We let $S^{(m)}$ be the set of points $p \in S$ for which $\underline{l}(p) \leq m$. Clearly $S^{(m)} \subset S^{(m+1)} \subset S^{(m+2)} \ldots$ and

$$\lim_{m\to\infty} S^{(m)}=S.$$

We take some $\underline{m}$ and take as the sequence in the lemma a sequence of $n + m$ numbers (we use $n + m$ in place of $\underline{n}$)

$$\varphi(p)-\beta;\ \varphi(Tp)-\beta;\ \ldots\ \varphi(T^{n+m-1}p)-\beta.$$

We now take for every $p \in S$ the sum of the selected numbers

$$\varepsilon_1(p)(\varphi(p)-\beta)+\varepsilon_2(p)(\varphi(Tp)-\beta)+\ldots+\varepsilon_{n+m}(p)(\varphi(T^{n+m-1}p)-\beta)>0,$$

where $\varepsilon_i(p)$ is either 0 or 1 ($i = 1, 2, \ldots, n + m$). Integrating over the set S, we find

$$\sum_{k=1}^{n+m}\int_{S_k}(\varphi(T^{k-1}p)-\beta)\,d\mu>0,$$

where S_k designates the set of those $p \in S$ for which $\varphi(T^{k-1}p) - \beta$ is a selected number. We note that $S_1 = S^{(m)}$. Moreover, when $k \leq n$, $TS_k = S_{k-1}$ and $S_k = T^{-1}S_{k-1}$ ($k \geq 2$). In fact, let $p \in S_k$; then there is an $\underline{l} \leq m$ such that

$$\varphi(T^{k-1}p)-\beta+\varphi(T^{k}p)-\beta+\ldots+\varphi(T^{k+l-1}p)-\beta>0.$$

but

$$\begin{aligned}0<\varphi(T^{k-1}p)-\beta+\varphi(T^{k}p)-\beta+\ldots+\varphi(T^{k+l-1}p)-\beta=\\=\varphi(T^{k-2}Tp)-\beta+\varphi(T^{k-1}Tp)-\beta+\ldots+\varphi(T^{k-2+l}Tp)-\beta.\end{aligned}$$

Consequently $Tp \in S_{k-1}$ (we do not leave S, since S is invariant).

Thus $TS_k \subset S_{k-1}$. Moreover, let $p' \in S_{k-1}$. We take any preimage $\underline{p}$ of point p': $p' = Tp$, $p \in T^{-1}S_{k-1}$.

$$\varphi(T^{k-2}p')+\varphi(T^{k-1}p')+\ldots+\varphi(T^{k-2+l'}p')>0,\ l'<m.$$

But since $p' = Tp$ this expression will equal

$$\varphi(T^{k-1}p)+\varphi(T^kp)+\ldots+\varphi(T^{k-1+l^*}p),$$

i.e., $p \in S_k$. Thus $T^{-1}S_{k-1} \subset S_k$. This, together with $TS_k \subset S_{k-1}$, proves the statement.

In view of this statement and the invariance of the measure μ we obtain

$$\sum_{k=1}^{n+m} \int_{S_k} (\varphi(T^{k-1}p)-\beta)\,d\mu = n \int_{S^{(m)}} (\varphi(p)-\beta)\,d\mu +$$

$$+\sum_{k=n+1}^{n+m} \int_{S_k} (\varphi(T^{k-1}p)-\beta)\,d\mu.$$

We may write (in view of $\int_{S_k} (\varphi(T^{k-1}p)-\beta)\,d\mu < \int_{S} |\varphi(p)-\beta|\,d\mu$) the inequality

$$\int_{S^{(m)}} (\varphi(p)-\beta)d\mu + \frac{m}{n}\int_{S} |\varphi(p)-\beta|\,d\mu > 0.$$

Thus (as $\underline{n}$ approaches infinity)

$$\int_{S^{(m)}} (\varphi(p)-\beta)\,d\mu > 0.$$

But since $\lim_{m\to\infty} S^{(m)} = S$,

$$\int_{S} (\varphi(p)-\beta)\,d\mu > 0,$$

i.e.,

$$\int_{S} \varphi(p)\,d\mu > \beta\mu S.$$

Moreover, if $p \in S$, there exists an $\underline{n}$ such that

$$\frac{\varphi(p)+\varphi(Tp)+\ldots+\varphi(T^{n-1}p)}{n} < \alpha$$

or

$$\alpha-\varphi(p)+\alpha-\varphi(Tp)+\ldots+\alpha-\varphi(T^{n-1}p) > 0.$$

Repeating the entire argument we obtain

$$\int_{S} (\alpha-\varphi(p))\,d\mu \geq 0, \quad \alpha\mu S \geq \int_{S} \varphi(p)\,d\mu,$$

i.e., $\alpha\mu S \geq \beta\mu S$.

Since $\mu S > 0$, $\alpha \geq \beta$, which contradicts the assumption that $\beta > \alpha$. The theorem is proved.

We let

$$\psi(p) = \lim_{n\to\infty} \frac{\varphi(p)+\varphi(Tp)+\ldots+\varphi(T^{n-1}p)}{n}.$$

The function $\psi(p)$ is defined nearly everywhere on R. We note that the function $\psi(p)$ is absolutely summable with respect to the measure μ. In fact,

$$\int_R \left|\frac{1}{n}\sum_{k=0}^{n-1}\varphi(T^k p)\right| d\mu < \frac{1}{n}\int_R \sum_{k=1}^{n} |\varphi(T^k p)|\, d\mu = \int_R |\varphi(p)|\, d\mu$$

(this last by the invariance of the measure). From this and from Fatou's theorem ([40], page 155) it follows that $\int_R |\psi(p)|\, d\mu$ exists. From this we conclude that $\psi(p)$ is finite almost everywhere.

Let us prove a lemma.

<u>Lemma. For any $\varepsilon > 0$, a δ can be found such that for any set A with $\mu A < \delta$ and for any $n \geq 1$</u>

$$\int_A \left|\frac{1}{n}\sum_{k=0}^{n-1}\varphi(T^k p)\right| d\mu < \varepsilon.$$

Proof.

$$\int_A \left|\frac{1}{n}\sum_{k=0}^{n-1}\varphi(T^{k-1} p)\right| d\mu < \frac{1}{n}\sum_{k=1}^{n-1}\int_A |\varphi(T^k p)|\, d\mu.$$

Let $T^{-k}A$ be the complete kth-order preimage of set A. Making a change in variables, we obtain

$$\int_A \left|\frac{1}{n}\sum_{k=1}^{n}\varphi(T^k p)\right| d\mu < \frac{1}{n}\sum_{k=1}^{n-1}\int_{T^{-k}A} |\varphi(p)|\, d\mu.$$

$\mu T^{-k}A = \mu A$. Since the function $\varphi(p)$ is absolutely summable, for $\varepsilon > 0$ there exists a $\delta > 0$ such that as soon as $\mu B < \delta$, $\int_B |\varphi(p)|\, d\mu < \varepsilon$. Making use of this fact, we find that

$$\int_A \left|\frac{1}{n}\sum_{k=1}^{n-1}\varphi(T^k p)\right| d\mu < \frac{1}{n}\cdot n\varepsilon = \varepsilon.$$

The lemma is proven. We also note that by Fatou's theorem, it will also be true that

$$\int_A |\psi(p)|\, d\mu < \varepsilon.$$

Let us prove the last assertion of the theorem. We have

$$\psi(p) = \lim_{n\to\infty} \frac{\varphi(p) + \ldots + \varphi(T^{n-1}p)}{n}.$$

We set $\varepsilon > 0$ and for the number ε and the function $\varphi(p)$ we select a δ in accordance with the lemma. On the basis of the Lebesgue theorem ([40], page 106) for the numbers $\varepsilon/3$ and δ there exists a $n_0(\varepsilon/3, \delta)$ such that for $n > n_0$

$$\mu E_p\left(\left|\psi(p) - \frac{1}{n}\sum_{k=0}^{n-1}\varphi(T^k p)\right| > \frac{\varepsilon}{3}\right) < \delta.$$

For $n > n_0$ we evaluate the difference:

$$\left|\int_R \psi(p)\,d\mu - \int_R \frac{1}{n}\left(\sum_{k=0}^{n-1}\varphi(T^k p)\right)d\mu\right| <$$

$$< \int_R \left|\psi(p) - \frac{1}{n}\sum_{k=0}^{n-1}\psi(T^k p)\right| d\mu <$$

$$< \int_{R-E_p} \left|\psi(p) - \frac{1}{n}\sum_{k=0}^{n}\varphi(T^k p)\right| d\mu + \int_{E_p} |\psi(p)|\,d\mu +$$

$$+ \int_{E_p} \left|\frac{1}{n}\sum_{k=0}^{n-1}\varphi(T^k p)\right| d\mu < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

(the first inequality is obtained from the construction of E_p, the second is a consequence of the lemma, and the third is obtained from the lemma). Thus

$$\int_R \psi(p)\,d\mu = \lim_{n\to\infty}\int_R \frac{1}{n}\left(\sum_{k=0}^{n-1}\varphi(T^k p)\right)d\mu.$$

but

$$\int_R \frac{1}{n}\left(\sum_{k=0}^{n-1}\varphi(T^k p)\right)d\mu = \frac{1}{n}\sum_{k=0}^{n-1}\int_R \varphi(T^k p)\,d\mu.$$

Since the complete preimage (of any order) of R is R itself

$$\int_R \varphi(T^k p)\,d\mu = \int_R \varphi(p)\,d\mu.$$

Hence

$$\int_R \psi(p)\,d\mu = \int_R \varphi(p)\,d\mu.$$

The theorem is proved.

Theorem (part 2). If a dynamic system is indecomposable with respect to a measure μ then for nearly all points $p \in R$ (with respect to the measure μ)

$$\lim_{n\to\infty} \frac{\varphi(p)+\varphi(Tp)+\ldots+\varphi(T^{n-1}p)}{n} = \int_R \varphi(p)\,d\mu.$$

Proof. It is sufficient to show that $\psi(p)$ is constant almost everywhere. In fact, let $\psi(p) = c$ nearly everywhere; then by the first part of the theorem: $c\int_R d\mu = \int_R \varphi(p)\,d\mu,$

$$c\mu R = \int_R \varphi(p)\,d\mu.$$

But $\mu R = 1$. Thus $\psi(p) = \int_R \varphi(p)\,d\mu,$ which we also need.

We shall prove that $\psi(p)$ is constant nearly everywhere. We let M be the upper bound on the function $\psi(p)$ calculated with an accuracy of up to the set of measure zero, i.e., $\mu E_p(\psi(p) > M) = 0$, while for any $\varepsilon > 0$ $\mu Ep(\psi(p) > M - \varepsilon) > 0$, we let $\underline{m}$ be the lower bound on the function $\psi(p)$, calculated with an accuracy of up to the set of measure zero. We must show that $M = m$. We assume that $M \neq m$, i.e., let there be an α such that $m < \alpha < M$. By definition of M, $\mu E_p(\psi(p) \geq \alpha) > 0$ (the inequality is strict) and by the definition of $\underline{m}$

$$\mu(R \setminus E_p(\psi(p) > \alpha)) = \mu E_p(\psi(p) < \alpha) > 0.$$

But the set $E_p(\psi(p) \geq \alpha)$ and $E_p(\psi(p) < \alpha)$ is invariant. This follows from the fact that the sums

$$\frac{\varphi(p)+\varphi(Tp)+\ldots+\varphi(T^{n-1}p)}{n}$$

and

$$\frac{\varphi(Tp)+\varphi(TTp)+\ldots+\varphi(T^{n-1}Tp)}{n}$$

differ by an expression of the order of $O(1/n)$ for nearly all $\underline{p}$. Thus R has been decomposed into the sum of two invariant sets of positive measure, which contradicts the hypothesis.

§21. Proof of the Normality Criterion for Sequences of Symbols on the Basis of the Birkhoff Theorem

We shall now prove the normality of sequences of symbols by the following method, based on the Birkhoff theorem. We note that the idea underlying the proof of §4 also underlies this proof. Here the Birkhoff theorem plays a role similar to that played by Lemma 1, §4.

We shall prove a lemma similar to the Kelly lemma ([40], page 240).

Lemma 1. Let there be given some fixed sequence made up of the symbols 0, 1, ..., $g - 1$,

$$\alpha = \varepsilon_1\varepsilon_2 \ldots$$

There exists a sequence of natural numbers

$$n_1 < n_2 < n_3 \ldots$$

such that for any elementary cylinder set the limit

$$\lim_{i\to\infty} \frac{N_{n_i}(\Delta)}{n_i}.$$

exists.

Proof. We may label all of the elementary cylinder sets Δ_1, Δ_2, ...

The numbers $N_n(\Delta)/n$ lie on the segment [01]. Since these numbers form a bounded set, we can find a subsequence of numbers such that $\lim N_{n_{1i}}(\Delta_1)/n_{1i}$ exists; we call this limit $\bar{\mu}\Delta_1$. We select the convergent sequence $N_{n_{2i}}(\Delta_2)/n_{2i}$, whose limit we designate $\bar{\mu}\Delta_2$, from the set of numbers $N_{n_{1i}}(\Delta_2)/n_{2i}$). We note that $\lim\limits_{n_{2i}\to\infty} \frac{N_{n_{2i}}(\Delta_1)}{n_{2i}} = \bar{\mu}\Delta_1$. We continue to choose sequences $\{n_{1i}\}$ such that

$$\lim_{n_{li}\to\infty} \frac{N_{n_{li}}(\Delta_j)}{n_{li}} = \bar{\mu}\Delta_j, \quad j = 1, 2, \ldots, l.$$

We now put together a diagonal sequence of numbers $n_1 = n_{11}$, $n_2 = n_{22}$, ... It is clear that $\lim\limits_{n_i\to\infty} \frac{N_{n_i}(\Delta_j)}{n_i} = \bar{\mu}\Delta_j$ for any j.

We note that $\bar{\mu}R = 1$. We note further that we can extend the function $\bar{\mu}$ to all cylinder sets and that this function will possess

complete additivity. In short, we can introduce the measure $\bar{\mu}$ into R.

Let us now prove the criteria. Let the sequence

$$\alpha = a_1, a_2, a_3, \ldots \tag{1}$$

satisfy the hypothesis of the criterion.

We assume that this sequence is not normal. This means that there exists an elementary cylinder set $\mathfrak{M}$ and a sequence of integers n_1, n_2, ..., n_k, ... and a number $\varepsilon > 0$ such that

$$\left| \frac{N_{n_k}(\mathfrak{M})}{n_k} - \mu\mathfrak{M} \right| > \varepsilon.$$

By Lemma 1 of this section, there exists a sequence n_{k_i} such that for any cylinder set Δ

$$\lim_{i\to\infty} \frac{N_{n_{k_i}}(\Delta)}{n_{k_i}} = \mu\Delta.$$

exists. Thus,

$$\frac{N_X(\Delta)}{X} = \frac{N_X(T^{-1}\Delta)}{X} + \frac{2\theta}{X}, \quad |\theta| < 1,$$

$$\bar{\mu}\Delta = \bar{\mu}T^{-1}\Delta,$$

i.e., $\bar{\mu}\Delta$ is an invariant normalized measure in a dynamic system.

We let $\mathfrak{M}^*$ be the set of all sequences p for which

$$\lim_{X\to\infty} \frac{N_X(p, \mathfrak{M})}{X} = \mu\mathfrak{M}.$$

By the Birkhoff theorem (part two) we obtain

$$\mu\mathfrak{M}^* = 1.$$

when we take the characteristic function of the set $\mathfrak{M}$ for $\varphi(p)$ and make use of the indecomposability with respect to the measure μ.

This has as its consequence

$$\lim_{X\to\infty} \frac{N_X(\alpha, \mathfrak{M}^*)}{X} = 1.$$

In fact, let $C\mathfrak{M}^*$ be the complement of $\mathfrak{M}^*$ in R. This will be the measurable set $\mu C\mathfrak{M}^* = 0$, since as is required by the criterion

$$0 < \overline{\lim_{X\to\infty}} \frac{N_X(\alpha, C\mathfrak{M}^*)}{X} < C\mu C\mathfrak{M}^* = 0.$$

We once again make use of the Birkhoff theorem. The lemma

$$\lim_{X\to\infty} \frac{N_X(p, \mathfrak{M})}{X} = \phi(p).$$

exists for almost all p with respect to the measure $\overline{\mu}$. But

$$\lim_{X\to\infty} \frac{N_X(\alpha, \mathfrak{M}^*)}{X} = 1$$

and this means that

$$\lim_{X\to\infty} \frac{N_X(p, \mathfrak{M})}{X} = \mu\mathfrak{M}.$$

for almost all p with respect to the measure $\overline{\mu}$.

We make use of the Lebesgue theorem ([40], pages 106 and 108) (if a functional sequence converges nearly everywhere with respect to the measure $\overline{\mu}$, it will then also converge with respect to the measure $\overline{\mu}$), applying it to the sequence of functions

$$\frac{\chi(p)+\chi(Tp)+\ldots+\chi(T^n p)}{n} = \frac{N_n(p, \mathfrak{M})}{n}, \quad n=1, 2, \ldots,$$

where $\chi(p)$ is a characteristic function of the set $\mathfrak{M}$. Applying another of the Lebesgue theorems ([40], page 139) to this functional sequence and remembering that $\frac{N_X(p, \mathfrak{M})}{X} < 1$ and $\int_R 1 d\overline{\mu} = 1$, we conclude that:

$$\lim_{X\to\infty} \int_R \frac{N_X(p, \mathfrak{M})}{X} d\overline{\mu} = \int_R \mu\mathfrak{M} d\overline{\mu} = \mu\mathfrak{M}.$$

But by the invariance of $\overline{\mu}$

$$\frac{1}{X}\int_R N_X(p, \mathfrak{M}) d\overline{\mu} = \int_R N_1(p, \mathfrak{M}) d\overline{\mu} = \overline{\mu}\mathfrak{M}.$$

Thus, $\overline{\mu}\mathfrak{M} = \mu\mathfrak{M}$.

In particular, this means that

$$\lim_{n_{k_l}\to\infty} \frac{N_{n_{k_l}}(\mathfrak{M})}{n_{k_l}} = \mu\mathfrak{M}.$$

But

$$\left| \frac{N_{n_{k_l}}(\mathfrak{M})}{n_{k_l}} - \mu\mathfrak{M} \right| > \varepsilon.$$

The contradiction proves the criterion.

§22. Dynamic System Corresponding to the Simplest Markov Chain

Let there be given two nonnegative numbers p_0 and p_1 which we shall call "initial probabilities," and four nonnegative numbers, which we shall call transition probabilities

$$\begin{pmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{pmatrix},$$

where

$$p_{00}+p_{01}=1,$$
$$p_{10}+p_{11}=1.$$

We shall assume that the stationarity conditions

$$p_{10}=p_0p_{01}+p_0p_{10},$$
$$p_{01}=p_1p_{01}+p_1p_{10}.$$

are valid (it is clear from this that $p_0 + p_1 = 1$).

We shall construct a dynamic system similar to that constructed by us in §19.

The space R is a space of infinite sequences made up of the symbols 0 and 1. We shall call an elementary cylinder set a subset of R if some number of the first symbols among its elements are fixed. We shall use $M(\delta_1 \ldots \delta_s)$ to designate an elementary cylinder set.

We introduce a function μ defined on elementary cylinder sets

1. $\mu R=1, \quad \mu 0=0.$
2. $\mu M(\delta_1\delta_2 \ldots \delta_s)=p_{\delta_1}p_{\delta_1\delta_2}\cdots p_{\delta_{s-1}\delta_s}.$

As in §18, we introduce the notion of a cylinder set and prove that cylinder sets form an algebra. In the analog of Lemma 2 of §18, the proof of the third point is more cumbersome: those $\gamma_1 \ldots \gamma_N$ generated in M intersect the $M_1(\delta_1^{(1)} \ldots \delta_{s_1}^{(1)})$. They are elementary cylinder sets of the form $\delta_1^{(1)} \ldots \delta_{s_1}^{(1)}\tau_1 \ldots \tau_{s-s_1}$

$$\sum_{\tau_1=0}^{1}\sum_{\tau_2=0}^{1}\cdots\sum_{\tau_{s-s_1}=0}^{1} p_{\delta_1^{(1)}}p_{\delta_1^{(1)}\delta_2^{(1)}}\cdots p_{\delta_{s_1}^{(1)}\tau_1}p_{\tau_1\tau_2}\cdots p_{\tau_{s-s_1-1}\tau_{s-s_1}}=$$

$$=p_{\delta_1^{(l)}}p_{\delta_1^{(l)}\delta_2^{(l)}}\cdots\sum_{\varepsilon_1=0}^{1}p_{\delta_{s_l}^{(l)}\varepsilon_1}\cdots\sum_{\varepsilon_{s-s_l}=0}^{1}p_{\varepsilon_{s-s_l-1}\varepsilon_{s-s_l}}=$$

$$=p_{\delta_1^{(l)}}p_{\delta_1^{(l)}\delta_2^{(l)}}\cdots p_{\delta_{s_l-1}^{(l)}\delta_{s_l}^{(l)}}=\mu M(\delta_1^{(l)}\dots\delta_{s_l}^{(l)})$$

(since $\sum\limits_{\varepsilon_{s-s_l}=0}^{1}p_{\varepsilon_{s-s_l-1}\varepsilon_{s-s_l}}=1$).

We now construct the measure μ on the minimum σ algebra that contains the class of cylinder sets.

We define the family of transformations T^k, $k = 0, 1, \dots$ in R. If $p = \varepsilon_1\varepsilon_2\dots$, then

$$T^k p=\varepsilon_{k+1}\varepsilon_{k+2}\dots$$

Clearly

$$T^{k_1+k_2}p=T^{k_1}(T^{k_2}p)$$

for any integers k_1 and $k_2 \geq 0$. The transformation T is not one-to-one.

Lemma 1. The measure μ is an invariant measure.

Proof. The complete preimage of the set $M(\delta_1\dots\delta_s)$ consists of $M_1(0\delta_1\dots\delta_s)$ and $M_2(1\delta_1\dots\delta_s)$. M_1 and M_2 are disjoint

$$\mu M_1+\mu M_2=p_0p_{0\delta_1}p_{\delta_1\delta_2}\dots p_{\delta_{s-1}\delta_s}+p_1p_{1\delta_1}p_{\delta_1\delta_2}\dots p_{\delta_{s-1}\delta_s}=$$

$$=p_{\delta_1\delta_2}\dots p_{\delta_{s-1}\delta_s}(p_0p_{0\delta_1}+p_1p_{1\delta_1})=p_{\delta_1\delta_2}\dots p_{\delta_{s-1}\delta_s}\left(\frac{p_{10}p_{0\delta_1}}{p_{10}+p_{01}}+\frac{p_{01}p_{1\delta_1}}{p_{10}+p_{01}}\right)=$$

$$=p_{\delta_1}p_{\delta_1\delta_2}\dots p_{\delta_{s-1}\delta_s}=\mu M$$

(in view of the fact that $p_{10}p_{0\delta_1} + p_{01}p_{1\delta_1} = p_{1-\delta_1\,\delta_1}$).

Lemma 2. Let the elements of the matrix $\begin{pmatrix}p_{00} & p_{01}\\ p_{10} & p_{11}\end{pmatrix}$ be positive. The space R cannot be represented as the sum of two invariant disjoint sets, both with positive measure, $R = U_1 \cup U_2$.

Proof. Let us assume that we are able to represent R as the sum of two invariant sets with positive measure, not having common points, $R = U_1 \cup U_2$. We let $\eta = \mu U_1$, $0 < \eta < 1$. Let $\chi(p)$ be the characteristic function of set U_1. Let there be a combination $\varepsilon_1\varepsilon_2\dots\varepsilon_n$ and let $\alpha = \delta_1\delta_2\dots$ be some point in R. Since the sequence

$$\varepsilon_1\varepsilon_2\ldots\varepsilon_n\delta_1\delta_2\ldots$$

is one of the nth-order preimages of point p, and U_1 is an invariant set,

$$\chi(\varepsilon_1\varepsilon_2\ldots\varepsilon_n\alpha)=\chi(\alpha).$$

The measure of the intersection of U_1 and the elementary cylinder set $M(\varepsilon_1\varepsilon_2\ldots\varepsilon_n)$ will equal

$$\int\limits_{M(\varepsilon_1\varepsilon_2\ldots\varepsilon_n)}\chi(\alpha)\,d\mu.$$

Let T' be a transformation mapping R into $M(\varepsilon_1\varepsilon_2\ldots\varepsilon_n)$. Carrying out the inverse transformation we obtain

$$\int\limits_{M(\varepsilon_1\varepsilon_2\ldots\varepsilon_n)}\chi(\alpha)\,d\mu=\mu M(\varepsilon_1\varepsilon_2\ldots\varepsilon_n)\int\limits_R\chi(\varepsilon_1\varepsilon_2\ldots\varepsilon_n\alpha)\frac{p_{\varepsilon_n\delta_1}}{p_{\delta_1}}\,d\mu,$$

where δ_1 is the first symbol of α. Thus

$$\int\limits_{M(\varepsilon_1\varepsilon_2\ldots\varepsilon_n)}\chi(\alpha)\,d\mu=\mu M\int\limits_R\chi(\alpha)\frac{p_{\varepsilon_n\delta_1}}{p_{\delta_1}}\,d\mu=$$

$$=\mu M\left(\frac{p_{\varepsilon_n0}}{p_0}\int\limits_{M(0)}\chi(\alpha)\,d\mu+\frac{p_{\varepsilon_n1}}{p_1}\int\limits_{M(1)}\chi(\alpha)\,d\mu\right).$$

We let

$$\eta_1=\int\limits_{M(0)}\chi(\alpha)\,d\mu,\quad \eta_2=\int\limits_{M(1)}\chi(\alpha)\,d\mu.$$

η_1 and η_2 are independent of n and $M(\varepsilon_1\ldots\varepsilon_n)$

$$\eta_1<\int\limits_{M(0)}d\mu=p_0,\quad \eta_2<\int\limits_{M(1)}d\mu=p_1.$$

Since $\eta_1 + \eta_2 = \eta < p_0 + p_1 = 1$, one of the inequalities $\eta_1 \leq p_0$ and $\eta_2 \leq p_1$ is stripped. Thus

$$\frac{p_{\varepsilon_n0}}{p_0}\eta_1+\frac{p_{\varepsilon_n1}}{p_1}\eta_2<p_{\varepsilon_n0}+p_{\varepsilon_n1}=1$$

(the inequality is stripped). Thus

$$\mu(U_1\cap M(\varepsilon_1\ldots\varepsilon_n))=\mu M(\varepsilon_1\ldots\varepsilon_n)\eta',$$

where η' is strictly less than some $\rho < 1$ and is independent of n and $M(\varepsilon_1\ldots\varepsilon_n)$. By the hypothesis of the lemma $M(\varepsilon_1\ldots\varepsilon_n) \neq 0$.

Thus

$$\frac{\mu(U_1 \cap M(\varepsilon_1 \ldots \varepsilon_n))}{\mu(M(\varepsilon_1 \ldots \varepsilon_n))} < \rho < 1.$$

We assume that $\varepsilon > 0$ and $1 - \rho > \varepsilon$. In analogy with the accumulation-point theorem, the set U_1, being a set of positive measure, should have an accumulation point ϑ_0,

$$\vartheta_0 = \tau_1\tau_2 \ldots,$$

i.e., for any $\varepsilon > 0$ it is possible to find a δ_0 such that for any elementary cylinder set Δ containing ϑ_0 for which $\mu\Delta \leq \delta_0$

$$\frac{\mu(U_1 \cap \Delta)}{\mu\Delta} > 1 - \varepsilon$$

(here we make use of the restriction that the matrix elements be positive). We choose an n so large that for any cylinder set $M = (\kappa_1 \ldots \kappa_n)$, $\mu M \leq \delta_0$, and we take the set $\Delta = M(\tau_1\tau_2 \ldots \tau_n)$ for Δ. We then have

$$\mu(U_1 \cap \Delta) < \rho\mu\Delta$$

and

$$\mu(U_1 \cap \Delta) > (1 - \varepsilon)\mu\Delta.$$

This yields $\rho > 1 - \varepsilon$, which contradicts the hypothesis. This means that our assumption was erroneous, and the lemma is proved.

§23. Markov-Normal Sequences

Let s be a natural number. Consider any s-tuple made up of the symbols 0 and 1, $\Delta = (\delta_1\delta_2 \ldots \delta_s)$. The quantity $p_{\delta_1}p_{\delta_1\delta_2} \ldots p_{\delta_{s-1}\delta_s}$ will be designated as $\mu\Delta$.

Let there be an infinite sequence made up of symbols 0 and 1

$$\alpha = \varepsilon_1\varepsilon_2\varepsilon_3 \ldots \tag{1}$$

As previously, we form for any natural number s the sequence of tuples

$$(\varepsilon_1\varepsilon_2 \ldots \varepsilon_s)(\varepsilon_2\varepsilon_3 \ldots \varepsilon_{s+1}) \ldots (\varepsilon_P \ldots \varepsilon_{P+s-1}). \tag{2}$$

We let $N_P(\alpha, \Delta)$, or simply $N_P(\alpha)$, stand for the number of times a combination Δ appears prior to the Pth element of Sequence (2).

Definition. We call the sequence of symbols (1) Markov-normal if for any natural number s and any combination Δ_s of s terms the limit

$$\lim_{p\to\infty}\frac{N_p(\Delta_s)}{p}=\mu\Delta_s$$

exists.

Theorem. Let there be a sequence

$$a_1a_2a_3\ldots \tag{3}$$

such that there exists a constant C such that for any s and any combination Δ_s

$$\overline{\lim_{p\to\infty}}\frac{N_p(\Delta_s)}{p}<C\mu\Delta_s.$$

Then Sequence (3) is Markov-normal.

The rest of the proof follows without changing the proof of the theorem in §21.

§24. Construction of Markov-Normal Sequence of Symbols

The construction of a Markov-normal sequence of symbols given here is taken from [21].

Let there be a sequence of integers

$$2<\beta_1<\beta_2\ldots,$$

for which

$$\frac{\beta_r}{\beta_{r-1}}=1+O\left(\frac{1}{r}\right),$$

and for sequences of positive integers $\alpha_{00}^{(r)}$, $\alpha_{01}^{(r)}$, $\alpha_{10}^{(r)}$, $\alpha_{11}^{(r)}$, for which the relationship

$$\alpha_{00}^{(r)}+\alpha_{01}^{(r)}=\beta_r,$$
$$\alpha_{10}^{(r)}+\alpha_{11}^{(r)}=\beta_r,$$

holds, while when $r\to\infty$,

$$\frac{\alpha_{10}^{(r)}}{\beta_r}\to p_{10},\quad \frac{\alpha_{01}^{(r)}}{\beta_r}\to p_{01}.$$

For each r, we consider the auxiliary Markov chains having

transition probabilities

$$\begin{pmatrix} \frac{\alpha_{00}^{(r)}}{\beta_r} & \frac{\alpha_{01}^{(r)}}{\beta_r} \\ \frac{\alpha_{10}^{(r)}}{\beta_r} & \frac{\alpha_{11}^{(r)}}{\beta_r} \end{pmatrix}$$

and initial probabilities that satisfy the conditions

$$\alpha_{10}^{(r)} = \gamma_0^{(r)}\alpha_{10}^{(r)} + \gamma_0^{(r)}\alpha_{01}^{(r)},$$

$$\alpha_{01}^{(r)} = \gamma_1^{(r)}\alpha_{10}^{(r)} + \gamma_1^{(r)}\alpha_{01}^{(r)}.$$

We let $\mu^{(r)}$ be the corresponding measure. For any combination Δ

$$\mu^{(r)}\Delta = \mu\Delta + o(1),\ r \to \infty.$$

We let s_r be the sequence of all r-digit numbers in the binary system (including numbers beginning with zero); each number Δ will be repeated $(\alpha_{10}^{(r)} + \alpha_{01}^{(r)}\beta_r^{r-1}\mu^{(r)}\Delta)$ times. When the r-digit numbers are written, we shall separate them by apostrophes.

Theorem. <u>The sequence of symbols written symbolically as</u>

$$\alpha = s_1 s_2 s_3 \ldots,$$

<u>is Markov-normal</u>.

Proof. It is sufficient to show that there exists a constant C that for any $s \geq 1$ and any combination $\Delta_s = (\delta_1\delta_2 \ldots \delta_s)$ of <u>s</u> terms

$$\overline{\lim_{p \to \infty}} \frac{N_p(\Delta_s)}{p} < C\mu\Delta_s.$$

We make the following definitions:

x_r is the number of symbols 0 or 1 in s_r;

S_r is the sequence $s_1 s_2 \ldots s_r$;

X_r is the number of symbols 0 or 1 in Sr, $X_r = \sum_{j=1}^{r} x_j$

g_r is the number of occurrences of Δ_s in s_r;

G_r is the number of occurrences of Δ_s in S_r.

Let us evaluate s_r. We know that each r-digit number contains <u>r</u> symbols and is repeated $(\alpha_{10}^{(r)} + \alpha_{01}^{(r)})\beta_r^{r-1}\mu^{(r)}\Delta_r$ times. This means that

$$x_r = \sum_{\Delta_r}{}' r\,(\alpha_{10}^{(r)}+\alpha_{01}^{(r)})\,\beta_r^{r-1}\mu^{(r)}\Delta_r = r\,(\alpha_{10}^{(r)}+\alpha_{01}^{(r)})\,\beta_r^{r-1}$$

(the sum is taken over all combinations Δ_r and thus $\sum_{\Delta_r}{}' \mu^{(r)}\Delta_r = 1$).

We let $E\begin{pmatrix} i_1\, i_2 \ldots i_\lambda \\ \varepsilon_{i_1}\varepsilon_{i_2}\ldots\varepsilon_{i_\lambda}\end{pmatrix}$ be the number of sequences in which we find ε_{i_1} in the $i_{\underline{1st}}$ place, ε_{i_2} in the $i_{\underline{2nd}}$ place, etc. Since $\mu^{(r)}$ is an invariant measure, when $i_1 > 1$

$$\mu^{(r)}E\begin{pmatrix} i_1-1\; i_2-1 \ldots i_\lambda-1 \\ \varepsilon_{i_1} \quad \varepsilon_{i_2} \quad \ldots \quad \varepsilon_{i_\lambda}\end{pmatrix} = \mu^{(r)}E\begin{pmatrix} i_1\, i_2 \ldots i_\lambda \\ \varepsilon_{i_1}\varepsilon_{i_2}\ldots\varepsilon_{i_\lambda}\end{pmatrix}.$$

A combination Δ_s may enter into s_r either separated or not separated by apostrophes. If $r < s$, Δ_s cannot enter into s_r unseparated. Where $r \geq s$, Δ_s enters into s_r undivided exactly $(r-s+1)\beta_r^{r-1}(\alpha_{10}^{(r)}+\alpha_{01}^{(r)})\mu^{(r)}\Delta_s$ times. In fact, there are exactly $r - s + 1$ ways in which Δ_s can occupy an unseparated position in an r-digit number: The first term Δ_s may coincide with the first, second, ... $(r - s + 1)\underline{th}$ symbol of the r-digit number. Fixing the $\underline{k\text{th}}$ position, we find that there exist in s_r

$$\mu^{(r)}E\begin{pmatrix} k\; k+1 \ldots k+s-1 \\ \delta_1\;\; \delta_2 \;\ldots\;\; \delta_{s-1}\end{pmatrix}(\alpha_{10}^{(r)}+\alpha_{01}^{(r)})\beta_r^{r-1} =$$
$$= (\alpha_{10}^{(r)}+\alpha_{01}^{(r)})\beta_r^{r-1}\mu^{(r)}E\begin{pmatrix} 1\,2 \ldots s \\ \delta_1\delta_2\ldots\delta_s\end{pmatrix}$$

r-digit numbers in which Δ_s can occupy this position.

Thus Δ_s is contained undivided in s_r exactly $(r-s+1)\beta_r^{r-1}(\alpha_{10}^{(r)}+\alpha_{01}^{(r)})\mu^{(r)}\Delta_s$ times, and s_r contains $(\alpha_{10}^{(r)}+\alpha_{01}^{(r)})\beta_r^{r-1}$ apostrophes. A given apostrophe cannot separate more than $\underline{s}$ different Δ_s. Thus Δ_s enters undivided into s_r no more than $O(s(\alpha_{10}^{(r)}+\alpha_{01}^{(r)})\beta_r^{r-1}) = O\left(\frac{x_r}{r}\right)$ times (since $\underline{r}$ increases, $\underline{s}$ may involve a constant).

Thus

$$g_r = (r-s+1)(\alpha_{10}^{(r)}+\alpha_{01}^{(r)})\beta_r^{r-1}\mu^{(r)}\Delta_s + o(x_r) = x_r\mu^{(r)}\Delta_s + o(x_r).$$

Since

$$\mu^{(r)}\Delta_s = \mu\Delta_s + o(1),$$

then

$$g_r = x_r \mu\Delta_s + o(x_r).$$

Moreover

$$G_r = \sum_{k=1}^{r} g_k + O(r); \quad X_r = \sum_{k=1}^{r} x_k.$$

We obtain

$$\lim_{r\to\infty} \frac{G_r}{X_r} = \mu\Delta_s.$$

Let $X_{r-1} \leq P < X_r$

$$N_P(\Delta_s) < G_r, \quad X_{r-1} > (r-1)(\alpha_{10}^{(r-1)} + \alpha_{01}^{(r-1)})\beta_{r-1}^{r-2}.$$

$$\frac{1}{P} < \frac{X_r}{X_{r-1}}\frac{1}{X_r} = \frac{X_{r-1} + r(\alpha_{10}^{(r)} + \alpha_{01}^{(r)})}{X_{r-1}}\frac{1}{X_r} <$$

$$< \left(1 + \frac{r}{r-1} \cdot \frac{\alpha_{10}^{(r)} + \alpha_{01}^{(r)}}{\alpha_{10}^{(r-1)} + \alpha_{01}^{(r-1)}} \cdot \frac{\beta_r^{r-1}}{\beta_{r-1}^{r-2}}\right)\frac{1}{X_r}.$$

Since $\frac{\beta_r}{\beta_{r-1}} = 1 + O\left(\frac{1}{r}\right)$ and $\frac{\alpha_{10}^{(r-1)}}{\beta_{r-1}} + \frac{\alpha_{01}^{(r-1)}}{\beta_{r-1}} \to p_{10} + p_{01} > 0$ (strictly!),

$$\frac{1}{P} < C\frac{1}{X_r},$$

$$\overline{\lim_{P\to\infty}} \frac{N_P(\Delta_s)}{P} < C \lim_{r\to\infty} \frac{G_r}{X_r} = C\mu\Delta_s.$$

The criterion is satisfied and the theorem proven.

§25. The Dynamic System in the Theory of Continued Fractions

We assume the well-known theory of continued fractions (see [44]).

We take for R the set of infinite sequences of natural numbers. Let

$$\beta = c_1 c_2 c_3 \ldots$$

be some infinite sequence of natural numbers. We compare the natural irrational number

$$\beta' = \cfrac{1}{c_1 + \cfrac{1}{c_2 + \cfrac{1}{c_3 + \ldots}}}$$

with this sequence. This number lies within the segment [01].

Let E be some set of sequences, and E' the corresponding set of irrational numbers on the interval [01]. We will let $E\begin{pmatrix} i_1 & i_2 & \ldots & i_s \\ a_{i_1} & a_{i_2} & \ldots & a_{i_s} \end{pmatrix}$ be

the set of sequences defined by the fact that their elements contain ε_{i_1} in the $i_{\underline{1st}}$ position, ε_{i_2}, ..., in the $i_{\underline{2nd}}$ position, and ε_{i_s} in the $i_{\underline{sth}}$ position. If $E\binom{1\,2\,\ldots\,s}{a_1 a_2 \ldots a_s}$, then E' consists of irrational numbers from the interval $\left(\frac{p_{s-1}+p_s}{q_{s-1}+q_s}, \frac{p_s}{q_s}\right)$, where p_{s-1}/q_{s-1} and p_s/q_s are the next-to-last and last convergents of

$$\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{\ddots}{+\cfrac{1}{a_s}}}}$$

[the fact that the interval is written as $\left(\frac{p_{s-1}+p_s}{q_{s-1}+q_s}, \frac{p_s}{q_s}\right)$ may lead to the incorrect assumption that it is always the case that $\frac{p_{s-1}+p_s}{q_{s-1}+q_s}<\frac{p_s}{q_s}$; actually, the direction of the inequality depends on the parity of $\underline{s}$].

We call the set E measurable if the set E' is measurable. We use mes E to denote the Lebesgue measure of set E'. We let μE stand for the number

$$\mu E=\frac{1}{\log 2}\int_{E'}\frac{dx}{1+x}.$$

Clearly,

$$\mu R=\frac{1}{\log 2}\int_0^1\frac{dx}{1+x}=1.$$

Moreover, it is clear that

$$\frac{1}{2\log 2}\operatorname{mes}E<\mu E<\frac{1}{\log 2}\operatorname{mes}E. \tag{1}$$

We define in R the family of transformations [T^k, k = 0, 1, 2, ...: if $p = c_1c_2c_3\ldots$, then $T^kp = c_{k+1}c_{k+2}c_{k+3}\ldots$.

Clearly

$$T^{k_1+k_2}p=T^{k_1}(T^{k_2}p),\ k_1>0,\ k_2>0.$$

We let $T^{-1}E$ be the complete preimage of set E for the transformation T.

Lemma 1. <u>The measure</u> μ <u>is an invariant measure in a dynamic</u>

system.

Proof. We need to prove that $\mu T^{-1}E = \mu E$.

It clearly is sufficient to show that this relationship is true of the set E for which the corresponding set E' is a set of irrational numbers from the segment [Oα] (Fig. 2). The complete preimage of the interval (Oα), as we can see from the drawing, consists of the infinite number of intervals

$$\left(\frac{1}{1+\alpha}\,1\right)\left(\frac{1}{2+\alpha}\,\frac{1}{2}\right)\ldots$$

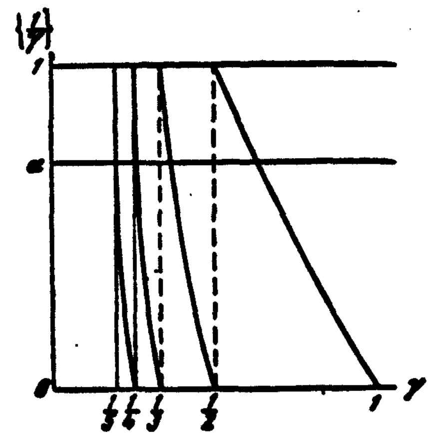

Fig. 2

The equation

$$\frac{1}{\log 2}\sum_{i=1}^{\infty}\int_{\frac{1}{i+\alpha}}^{\frac{1}{i}}\frac{dx}{1+x}=\frac{1}{\log 2}\sum_{i=1}^{\infty}\left[\log\left(1+\frac{1}{i}\right)-\log\left(1+\frac{1}{i+\alpha}\right)\right]=$$

$$=\frac{1}{\log 2}\lim_{P\to\infty}\left(\log P-\log\frac{P+\alpha}{1+\alpha}\right)=\frac{1}{\log 2}\lim_{P\to\infty}\log\frac{1+\alpha}{1+\frac{\alpha}{P}}=$$

$$=\frac{1}{\log 2}\log(1+\alpha)=\frac{1}{\log 2}\int_0^{\alpha}\frac{dx}{1+x}$$

proves the lemma.

A special case of this lemma is represented by the statement: when $i_1 > 1$

$$\mu E\begin{pmatrix} i_1 & i_2 & \ldots & i_s \\ \varepsilon_{i_1} & \varepsilon_{i_2} & \ldots & \varepsilon_{i_s}\end{pmatrix}=$$

$$=\mu E\begin{pmatrix} i_1-1 & i_2-1 & \ldots & i_s-1 \\ \varepsilon_{i_1} & \varepsilon_{i_2} & \ldots & \varepsilon_{i_s}\end{pmatrix}.$$

The lemma given below was proved by K. Knopp [45].

Lemma 2. The space R cannot be represented as the sum of two invariant sets with positive measure.

Proof. In view of Inequality (1) we need to establish that R cannot be represented as the sum of two invariant sets of positive Lebesgue measure. Let us assume that we are able to do this: $R = U_1 \cup \cup U_2$. Then the set of irrational numbers within the interval [01] is decomposed into the sum of two sets $U_1' \cup U_2'$. Let η = mes U_1'. The se U_1', being a set of positive measure, has an accumulation point

i.e., for $\varepsilon = (1-\eta)/(1+\eta) > 0$ there exists a δ such that if E' is an interval containing ϑ_0 with mes E < δ, then

$$\frac{\operatorname{mes}(E'\cap U_1')}{\operatorname{mes} E'} > 1-\varepsilon.$$

Let

$$\vartheta_0 = \cfrac{1}{c_1+\cfrac{1}{c_2+\ddots}}$$

We take for E' the interval consisting of numbers that have the first incomplete portions $c_1 c_2 \ldots c_n$ of length less than δ. Let this be the interval $\left(\frac{p_{n-1}+p_n}{q_{n-1}+q_n}\, \frac{p_n}{q_n}\right)$. We let $\chi(x)$ be the characteristic function of the set U_1'.

The number

$$y = \cfrac{1}{c_1+\cfrac{1}{c_2+\ddots+\cfrac{1}{c_n+x}}} = \frac{p_{n-1}x+p_n}{q_{n-1}x+q_n}$$

is one of the preimages of the number x. Thus

$$\chi\left(\frac{p_{n-1}x+p_n}{q_{n-1}x+q_n}\right) = \chi(x).$$

It is clear that $\operatorname{mes}(E'\cap U_1') = \int_{E'}\chi(x)\,dx$. $\operatorname{mes} E' = \left|\frac{p_{n-1}+p_n}{q_{n-1}+q_n}-\frac{p_n}{q_n}\right| = \frac{1}{(q_{n-1}+q_n)q_n}$. Thus

$$\frac{\operatorname{mes}(E'\cap U_1')}{\operatorname{mes} E'} =$$

$$= q_n(q_{n-1}+q_n)\int_{E'}\chi(x)\,dx = q_n(q_{n-1}+q_n)\int_0^1 \chi\left(\frac{p_{n-1}x+p_n}{q_{n-1}x+q_n}\right)\frac{dx}{(q_{n-1}x+q_n)^2} =$$

$$= q_n(q_{n-1}+q_n)\int_0^1 \chi(x)\frac{dx}{(q_{n-1}x+q_n)^2}.$$

Since $(q_{n-1}x+q_n)^2$ increases as x goes up, the interval can only increase if we assume that E' is the interval (0η).

$$\frac{\operatorname{mes}(E'\cap U_1')}{\operatorname{mes} E'} < q_n(q_{n-1}+q_n)\int_0^\eta \frac{dx}{(q_{n-1}x+q_n)^2} =$$

$$= 1-\frac{q_{n-1}(1-\eta)}{q_{n-1}\eta+q_n} < 1-\frac{1-\eta}{1+\eta}$$

which contradicts $\frac{\operatorname{mes}(E' \cap U_1')}{\operatorname{mes} E'} > 1 - \frac{1-\eta}{1+\eta}$. The lemma is proved.

§26. Normal Continued Fraction

Let s be a natural number. Consider any s-tuple consisting of natural numbers $\Delta = (\delta_1 \delta_2 \ldots \delta_s)$. Let Δ' be the set of all irrational numbers in the interval [01], and let the beginning of their decomposition into a continued fraction be

$$\cfrac{1}{\delta_1 + \cfrac{1}{\delta_2 + \cdots + \cfrac{1}{\delta_s}}}.$$

We define

$$\mu\Delta = \mu\Delta' = \frac{1}{\log 2}\int_{\Delta'} \frac{dx}{1+x}.$$

Let there be an infinite sequence consisting of the natural numbers

$$c_1 c_2 c_3 \ldots \tag{1}$$

We take any natural number s and write Sequence (1) as a caterpillar:

$$(c_1 c_2 \ldots c_s)(c_2 \ldots c_{s+1}) \ldots (c_p \ldots c_{p+s-1}) \ldots \tag{2}$$

We let $N_p(\Delta)$ denote the number of times the combination $\Delta = (\delta_1 \delta_2 \ldots \delta_s)$ is encountered prior to the Pth position of Sequence (2).

Definition. We call the sequence of natural numbers (1) a normal continued fraction if for any natural number s and any combination $\Delta = (\delta_1 \ldots \delta_s)$ of s terms composed of natural numbers the limiting relationship

$$\lim_{p\to\infty} \frac{N_p(\Delta)}{p} = \mu\Delta$$

holds.

We shall need the following theorem.

Theorem. Let there be the sequence

$$c_1 c_2 c_3 \ldots \tag{3}$$

such th[illegible] there exists a constant C such that for any s and any

combination $\Delta_s = (\delta_1 \ldots \delta_s)$ of s terms

$$\overline{\lim_{X \to \infty}} \frac{N_X(\Delta)}{X} < C \frac{1}{q_n(q_n + q_{n-1})}. \tag{4}$$

Then Sequence (3) is a normal continued fraction.

Proof. Let (4) be valid. Since

$$\frac{1}{q_n(q_n + q_{n-1})} < \frac{1}{2\log 2} \mu \Delta,$$

the relationship

$$\overline{\lim_{X \to \infty}} \frac{N_X(\Delta)}{X} < C' \mu \Delta.$$

holds.

We call the set of sequences whose first s terms are fixed an elementary cylinder set. The set of all elementary cylinder sets is denumerable. We can repeat the proof used for the theorem of §21.

§27. Construction of a Normal Continued Fraction

We present here the construction of a normal continued fraction due to A.G. Postnikov and I.I. Pyatetskiy [21].

We shall need the following notation:

Let s_1, $s_2 \ldots$ be groups of natural numbers

$$s_1 = c_{11} \ldots c_{1n_1}, \; s_2 = c_{21} \ldots c_{2n_2}, \ldots .$$

We let

$$a = s_1 s_2 \ldots$$

be a sequence of natural numbers

$$a = c_{11} \ldots c_{1n_1} c_{21} \ldots c_{2n_2}, \ldots .$$

We introduce the notation: $\underline{l}$, $\underline{r}$ are natural numbers $\underline{l} \to \infty$ and $r \to \infty$, but $r/\underline{l} \to 0$; $a_r^{(\underline{l})}$ is a row consisting of all r-digit groups $a_1 a_2 \ldots a_r$ where $1 \leq a_1 \leq \underline{l}, \ldots 1 \leq \underline{l} a_r \leq \underline{l}$($a_1, a_2, \ldots, a_r$ are natural numbers); here $a_1 a_2 \ldots a_r$ is repeated $\left[2^{r+\sqrt{r}}\, \mathrm{mes}\, E\begin{pmatrix} 1\,2 \ldots r \\ a_1 a_2 \ldots a_r \end{pmatrix}\right]$ times. These r-digit groups are separated from each other by apostrophes. The order of these groups within the row is unimportant.

We note that

$$2^{2r+1}l^{2r}\,\mathrm{mes}\,E\begin{pmatrix}1\,2\ldots r\\ a_1a_2\ldots a_r\end{pmatrix}>1,$$

i.e., the combination $a_1a_2\ldots a_r$ is actually found in $s_r^{(l)}$. As a matter of fact,

$$2^{2r+1}l^{2r}\,\mathrm{mes}\,E\begin{pmatrix}1\,2\ldots r\\ a_1a_2\ldots a_r\end{pmatrix}=\frac{2^{2r+1}l^{2r}}{q_r(q_r+q_{r-1})}>2^{2r+1}\frac{l^{2r}}{2q_r^2}>1,$$

since it is easily shown by induction that

$$q_r<(2l)^r,\quad q_{r+1}<lq_r+q_{r-1}<2lq_r<(2l)^{r+1}.$$

Moreover, we let $y_r^{(l)}$ be the number of groups in $s_r^{(l)}$.

It is clear that

$$y_r^{(l)}=\sum_{a_1=1}^{l}\cdots\sum_{a_r=1}^{l}\left[2^{2r+1}l^{2r}\,\mathrm{mes}\,E\begin{pmatrix}1\,2\ldots r\\ a_1a_2\ldots a_r\end{pmatrix}\right]=$$

$$=2^{2r+1}l^{2r}\sum_{a_1=1}^{l}\cdots\sum_{a_r=1}^{l}\mathrm{mes}\,E\begin{pmatrix}1\,2\ldots r\\ a_1a_2\ldots a_r\end{pmatrix}+O(l^r).$$

Let $U_r^{(l)}$ be the complement of the set

$$\bigcup_{a_1=1}^{l}\cdots\bigcup_{a_r=1}^{l}E\begin{pmatrix}1\,2\ldots r\\ a_1a_2\ldots a_r\end{pmatrix}$$

It is clear that

$$U_r^{(l)}=\bigcup_{j=l+1}^{\infty}E\begin{pmatrix}1\\ j\end{pmatrix}\cup\left(\bigcup_{l_1=1}^{l}\bigcup_{j=l+1}^{\infty}E\begin{pmatrix}1\,2\\ l_1j\end{pmatrix}\right)\cup\left(\bigcup_{l_1=1}^{l}\bigcup_{l_2=1}^{l}\bigcup_{j=l+1}^{\infty}E\begin{pmatrix}1\,2\,3\\ l_1l_2j\end{pmatrix}\right)$$

$$\cdots\cup\left[\bigcup_{l_1=1}^{l}\cdots\bigcup_{l_{r-1}=1}^{l}\bigcup_{j=l+1}^{\infty}E\begin{pmatrix}1\,2\ldots r-1\;r\\ l_1l_2\ldots l_{r-1}\;j\end{pmatrix}\right].$$

Since (see [44], page 78)

$$\mathrm{mes}\,E\begin{pmatrix}1\,2\ldots p\;p+1\\ l_1l_2\ldots l_p\;j\end{pmatrix}<C\,\mathrm{mes}\,E\begin{pmatrix}1\,2\ldots p\\ l_1l_2\ldots l_p\end{pmatrix}\frac{1}{j^2}$$

and, clearly,

$$\sum_{j=l+1}^{\infty}\frac{1}{j^2}=O\left(\frac{1}{l}\right),$$

then

$$\mathrm{mes}\,U_r^{(l)}=O\left(\frac{1}{l}\left(1+\sum_{l_1=1}^{l}\mathrm{mes}\,E\begin{pmatrix}1\\ l_1\end{pmatrix}+\sum_{l_1=1}^{l}\sum_{l_2=1}^{l}\mathrm{mes}\,E\begin{pmatrix}1\,2\\ l_1l_2\end{pmatrix}+\right.\right.$$

$$\left.\left.\cdots+\sum_{l_1=1}^{l}\cdots\sum_{l_r=1}^{l}\mathrm{mes}\,E\begin{pmatrix}1\ldots r\\ l_1\ldots l_r\end{pmatrix}\right)\right).$$

But

$$\sum_{l_1=1}^{l} \operatorname{mes} E\binom{1}{l_1} < 1, \sum_{l_1=1}^{l}\sum_{l_2=1}^{l} \operatorname{mes} E\binom{1\,2}{l_1 l_2} < 1, \ldots \operatorname{mes} U_r^{(l)} = O\left(\frac{r}{l}\right) = o(1),$$

since $r/\underline{l} \to 0$. Clearly

$$\sum_{a_1=1}^{l} \cdots \sum_{a_r=1}^{l} \operatorname{mes} E\binom{1 \ldots r}{a_1 \ldots a_r} + \operatorname{mes} U_r^{(l)} = 1.$$

Therefore

$$\sum_{a_1=1}^{l} \cdots \sum_{a_r=1}^{l} \operatorname{mes} E\binom{1 \ldots r}{a_1 \ldots a_r} = 1 + o(1).$$

We thus obtain

$$y_r^{(l)} = 2^{sr+1} l^{sr} + o(2^{sr} l^{sr}).$$

Let $x_r^{(\underline{l})}$ be the number of numbers in $s_r^{(\underline{l})}$. It is easy to see that $x_r^{(\underline{l})} = r y_r^{(\underline{l})}$, and it follows from this that

$$x_r^{(l)} = r 2^{sr+1} l^{sr} + o(r 2^{sr+1} l^{sr}).$$

Next let

$$\Delta = \Delta_\rho^\lambda = (a_1 \ldots a_\rho),$$

where $1 \leq a_j \leq \lambda$, $j = 1, 2, \ldots \rho$ (the top and bottom inequalities cannot be realized), λ and ρ are fixed.

We let $q_r^{(\underline{l})}$ be the number that shows how many times Δ is found in $s_r^{(\underline{l})}$. The combination Δ may be contained in $s_r^{(\underline{l})}$ separated or not separated by apostrophes; there are $r - \rho + 1$ ways in which Δ can be contained in any group without separation: the first term of Δ may coincide with the first, second, ..., $(r - \rho + 1)$th term of the group. Thus the quantity sought will equal

$$\sum_{a_{\rho+1}=1}^{l} \cdots \sum_{a_r=1}^{l} \left[2^{sr+1} l^{sr} \operatorname{mes} E\binom{1\,2 \ldots \rho\ \rho+1 \ldots r}{a_1 a_2 \ldots a_\rho\ a_{\rho+1} \ldots a_r}\right] +$$
$$+ \sum_{a_1=1}^{l}\sum_{a_{\rho+2}=1}^{l} \cdots \sum_{a_r=1}^{l} \left[2^{sr+1} l^{sr} \operatorname{mes} E\binom{1\,2 \ldots \rho+1\ \rho+2 \ldots r}{a_1 a_2 \ldots a_\rho\ \ a_{\rho+2} \ldots a_r}\right] +$$
$$+ \ldots + \sum_{a_1=1}^{l} \cdots \sum_{a_{r-\rho}=1}^{l} \left[2^{sr+1} l^{sr} \operatorname{mes} E\binom{1\,2 \ldots r-\rho\ r-\rho+1 \ldots r}{a_1 a_2 \ldots a_{r-\rho}\ \ a_1 \ldots\ \ a_\rho}\right] =$$

$$= 2^{\nu+1} l^{\nu} \Big(\sum_{a_{\rho+1}=1}^{l} \cdots \sum_{a_r=1}^{l} \operatorname{mes} E \begin{pmatrix} 1 & 2 \ldots \rho & \rho+1 \ldots r \\ a_1 & a_2 \ldots a_\rho & a_{\rho+1} \ldots a_r \end{pmatrix} + \ldots +$$

$$+ \sum_{a_1=1}^{l} \cdots \sum_{a_{r-\rho}=1}^{l} \operatorname{mes} E \begin{pmatrix} 1 \ldots r-\rho & r-\rho+1 \ldots r \\ a_1 \ldots a_{r-\rho} & a_1 \ldots \quad a_\rho \end{pmatrix} \Big) + O(rl^{r-\nu}) <$$

$$< 2^{\nu+1} l^{\nu} \Big(\sum_{a_{\rho+1}=1}^{\infty} \cdots \sum_{a_r=1}^{\infty} \operatorname{mes} E \begin{pmatrix} 1 & 2 \ldots \rho & \rho+1 \ldots r \\ a_1 & a_2 \ldots a_\rho & a_{\rho+1} \ldots a_r \end{pmatrix} +$$

$$\ldots + \sum_{a_1=1}^{\infty} \cdots \sum_{a_{r-\rho}=1}^{\infty} \operatorname{mes} E \begin{pmatrix} 1 \ldots r-\rho & r-\rho+1 \ldots r \\ a_1 \ldots a_{r-\rho} & a_1 \ldots \quad a_\rho \end{pmatrix} +$$

$$+ O(rl^{r-\nu}) = 2^{\nu+1} l^{\nu} \Big(\operatorname{mes} E \begin{pmatrix} 1 & 2 \ldots \rho \\ a_1 & a_2 \ldots a_\rho \end{pmatrix} + \operatorname{mes} E \begin{pmatrix} 2 & 3 \ldots \rho+1 \\ a_1 & a_2 \ldots a_\rho \end{pmatrix} +$$

$$+ \ldots + \operatorname{mes} E \begin{pmatrix} r-\rho+1 \ldots r \\ a_1 \quad \ldots \quad a_\rho \end{pmatrix} \Big) + O(rl^{r-\nu}) <$$

$$< C 2^{\nu+1} l^{\nu} \Big(\mu E \begin{pmatrix} 1 & 2 \ldots \rho \\ a_1 & a_2 \ldots a_\rho \end{pmatrix} + \ldots \mu E \begin{pmatrix} r-\rho+1 \ldots r \\ a_1 \quad \ldots a_\rho \end{pmatrix} \Big) + O(rl^{r-\nu}).$$

Because of the invariance of the measure

$$\mu E \begin{pmatrix} 1 & 2 \ldots \rho \\ a_1 & a_2 \ldots a_\rho \end{pmatrix} = \mu E \begin{pmatrix} 2 \ldots \rho+1 \\ a_1 \ldots \quad a_\rho \end{pmatrix} = \ldots = \mu E \begin{pmatrix} r-\rho+1 \ldots r \\ a_1 \quad \ldots a_\rho \end{pmatrix}.$$

Thus the quantity solved will be less than or equal to

$$C r 2^{\nu+1} l^{\nu} \mu E \begin{pmatrix} 1 & 2 \ldots \rho \\ a_1 & a_2 \ldots a_\rho \end{pmatrix} + O(rl^{r-\nu}) <$$

$$< C' r 2^{\nu+1} l^{\nu} \operatorname{mes} E \begin{pmatrix} 1 & 2 \ldots \rho \\ a_1 & a_2 \ldots a_\rho \end{pmatrix} + O(rl^{r-\nu}) <$$

$$< C'' x_r^{(l)} \operatorname{mes} E \begin{pmatrix} 1 & 2 \ldots \rho \\ a_1 & a_2 \ldots a_\rho \end{pmatrix} + o(x_r^{(l)}).$$

Let us determine how many times Δ may be contained divided in $s_r^{(\underline{l})}$. There are $y_r^{(\underline{l})} = x_r^{(\underline{l})}/r$ apostrophes in $s_r^{(\underline{l})}$. Any given apostrophe cannot separate more than ρ different groups. We shall thus have at most $\frac{x_r^{(l)}}{r} \cdot \rho = o(x_r^{(l)})$ possibilities.

Thus

$$g_r^{(l)} \leq C x_r^{(l)} \operatorname{mes} E \begin{pmatrix} 1 & 2 \ldots \rho \\ a_1 & a_2 \ldots a_\rho \end{pmatrix} + o(x_r^{(l)}).$$

Let $\underline{l} = 1, 2 \ldots$. We set $r = [\underline{l}n \, \underline{l}] + 1$. Clearly $r/\underline{l} \to 0$ when $\underline{l} \to \infty$. We denote $s_r^{(\underline{l})}$ simply by $s^{(\underline{l})}$ and consider the sequence

$$\alpha = s^{(1)} s^{(2)} s^{(3)} \ldots$$

We shall show that this is a normal continued fraction.

We introduce the notation:

$x^{(l)}$ is the number of symbols in $s^{(l)}$;

$g^{(l)}$ is the number of appearances of $\Delta = \Delta_{\rho}^{\lambda}$ in $s^{(l)}$;

$S^{(l)}$ is the row $s^{(1)}s^{(2)}\ldots s^{(l)}$;

$X^{(l)}$ is the number of terms in $S^{(l)}$.

Clearly

$$X^{(l)} = x^{(1)} + x^{(2)} + \ldots + x^{(l)}.$$

$G^{(l)}$ is the number of occurrences of Δ in $S^{(l)}$.

$$G^{(l)} = \sum_{k=1}^{l} g^{(k)} + O(l).$$

Since

$$g^{(l)} < Cx^{(l)} \operatorname{mes} E\begin{pmatrix} 1\,2\ldots\rho \\ \alpha_1\alpha_2\ldots\alpha_\rho \end{pmatrix} + o(x^{(l)}),$$

we obtain

$$G^{(l)} < CX^{(l)} \operatorname{mes} E\begin{pmatrix} 1\,2\ldots\rho \\ \alpha_1\alpha_2\ldots\alpha_\rho \end{pmatrix},$$

i.e.,

$$N_{X^{(l)}}(\Delta) < CX^{(l)} \operatorname{mes} E\begin{pmatrix} 1\,2\ldots\rho \\ \alpha_1\alpha_2\ldots\alpha_\rho \end{pmatrix} + o(X^{(l)}).$$

Let $X^{(l)} \leq P < X^{(l+1)}$. We have:

$$\frac{N_P(\Delta)}{P} < \frac{N_{X^{(l+1)}}(\Delta)}{X^{(l)}} < \frac{CX^{(l+1)} \operatorname{mes} E\begin{pmatrix} 1\,2\ldots\rho \\ \alpha_1\alpha_2\ldots\alpha_\rho \end{pmatrix} + o(X^{(l+1)})}{X^{(l)}}.$$

Since

$$\frac{X^{(l+1)}}{X^{l}} = \frac{X^{(l)} + x^{(l+1)}}{X^{(l)}} < 1 + \frac{x^{(l+1)}}{x^{(l)}}$$

and since

$$\frac{x^{(l+1)}}{x^{(l)}} = O\left(\frac{\ln(l+1)}{\ln l}\,\frac{2^{a\ln(l+1)}}{2^{a\ln l}}\,\frac{e^{a\ln^2(l+1)}}{e^{a\ln^2 l}}\right) = O(1),$$

then

$$\frac{X^{(l+1)}}{X^{(l)}} < C.$$

Thus

$$\varlimsup_{p \to \infty} \frac{N_p(\Delta)}{p} < C \operatorname{mes} E\begin{pmatrix} 1 & 2 & \ldots & p \\ a_1 & a_2 & \ldots & a_p \end{pmatrix}.$$

The criterion for a normal continued fraction is satisfied.

NOTES

[1] page 4. Reichenbach [46] deals with this subject.

Copeland [47] gives some idea of the importance of admissible numbers for problems in which the frequency is given geometrically. There is a summary of the application of random and pseudorandom numbers given in [48]. This class of problems is also referred to in [49].

[2] page 16. A.G. Postnikov [52] put Champernowne's construction into geometric form and constructed a complex number $\alpha + \beta i$ such that the fractions $\{(\alpha + \beta i)(a + bi)^x\}$, $x = 1, 2, 3\ldots$, are uniformly distributed (here $a + bi$ is the fixed Gaussian integer, differing from $-1, 1, i, -i$). A.M. Polosuyev [53] also developed a method for the case in which there are restrictions on an n-order integer matrix A, and constructed a vector $\bar{\alpha}$ such that the fractions $\{\bar{\alpha}A^x\}$, $x = 1, 2, 3\ldots$, are uniformly distributed in a unit cube in n-dimensional space.

Normal periodic systems have been used to obtain very good remainder terms in a problem dealing with the uniform distribution of fractional portions of an exponential function. N.M. Korobov [36] constructed a system of real numbers $\alpha_1, \ldots, \alpha_s$ such that for any s and any system of natural numbers larger than one, $g_1, \ldots, g_s$, the condition

$$N_P(\Delta) = P \operatorname{mes} \Delta + O\left(P^{1-\frac{1}{s+1}}\right)$$

(the constant in "O" depends on Δ) will hold for the number of

occurrences of point $\{(\alpha_1 g_1^x) \dots (\alpha_s g_s^x)\}$ within any parallelepiped Δ lying within a unit cube with sides parallel to the coordinate axes.

A.G. Postnikov [54] has constructed a real number α for which the condition

$$N_P(\Delta) = P \operatorname{mes} \Delta + O\left(\frac{\sqrt{P}}{\sqrt[s]{\log P}} \log \log P\right).$$

holds.

N.M. Korobov has solved several problems in which the problem of constructing numbers α for which fractions $\{\alpha g^x\}$, $x = 1, 2, \dots$, are uniformly distributed occurs as a special case [55-57]. The methods of proof in these papers are based on using evaluations of various types of trigonometric sums. A study of A.M. Polosuyev dealt with a generalization of one of the results [58].

We should take note of N.M. Korobov's study evaluating the sums of fractional portions of an exponential function [59]. Finally, N.M. Korobov has investigated a solution to inhomogeneous Diophantine inequalities with exponential functions [36].

[3] page 59 We shall apply the Birkhoff-Khinchin theorem to a dynamic system in a space of sequences of symbols. In this problem, we can reduce the case of transformations that are not one-to-one to transformations that are one-to-one, thus eliminating the need for using the Reiz theorem.

We compare with each infinity in the sequence of real numbers made up of the symbols 0 and 1,

$$\varepsilon_1 \varepsilon_2 \varepsilon_3 \dots$$

with the sequence, infinite in both directions,

$$\dots \varepsilon_{-2} \varepsilon_{-1} \varepsilon_0 \varepsilon_1 \varepsilon_2 \dots,$$

in which

$$\varepsilon_{-i} = \varepsilon_{i+1}, \quad i > 0.$$

We define a group of one-to-one transformations on the set of sequences of symbols, infinite in both directions, by means of the equation

$$T^k(\ldots s_{-1}\, s_0\, s_1 \ldots) = \ldots s_{k-1}\, s_k\, s_{k+1} \ldots,$$

k = ... -1, 0, 1, 2... . An investigation of dynamic systems in a space of sequences of symbols infinite in one direction is reduced to studying a dynamic system in a space of sequences of symbols infinite in both directions.

REFERENCES

1. N.M. Korobov, O nekotorykh voprosakh ravnomernogo raspredeleniya [Certain Uniform-Distribution Problems], Izv. Akad. nauk SSSR, ser. matem. [Bulletin of Acad. Sci. USSR, Math. Series], 14 (1950), 215-231.
2. H. Weyl, Ueber die Gleichverteilung von Zahlen mod. Eins [Uniform Distribution of Numbers Modulo One], Math. Annalen [Mathematical Annals], 77 (1916), 313-352.
3. E. Borel, Lecon sur la theorie des fonctions [Lesson on the Theory of Functions], Paris, 1914.
4. M.G. Kendall and B. Babington Smith, Randomness and Random Sampling Numbers, Journ. of the Royal Statistical Society, 101, 1, 1938, 147-166.
5. H. Steinhaus, Ueber einige prinzipielle Fragen der mathematischen Statistik, Ber. Tagung Wahrscheinlichkeitsrechnung und mathematischen Statistik [Certain Fundamental Problems of Mathematical Statistics. Reports of the Conference on Probability Theory and Mathematical Statistics], Berlin, 1954, Berlin, 1956.
6. I.M. Sobol', Psevdosluchaynyye chisla dlya mashiny "Strela." Teoriya veroyatnostey i yeye primeneniya [Pseudorandom Numbers for the "Strela" Computer. Probability Theory and its Applications], 3, No. 2 (1958), 205-211.
7. I.A. Venn, The Logic of Chance, London, 1888.
8. R. Mizes, Veroyatnost' i statistika [Probability and Statistics], Moscow, 1930.
9. R. Peter, Rekursivnyye funktsii [Recursive Functions], Moscow, 1954.
10. A. Church, On the Concept of a Random Sequence, Bull. of the A. Math. Soc., 46, 1940, 130-135.

11. A.H. Copeland, Admissible Numbers in the Theory of Probability, Amer. Journ. of Math., 50, 1928, pp. 535-552.

12. A.G. Postnikov and I.I. Pyatetskiy, Normal'nyye po Bernulli posledovatel'nosti znakov [Normal Number Sequences According to Bernoulli], Izv. AN SSSR, ser. matem. [Bulletin of the Acad. Sci. USSR, Math. Series, 21 (1957), 501-514.

13. D.L. Dub, Veroyatnostnyye protsessy [Probability Processes], Moscow, 1956.

14. A.G. Postnikov, Resheniye sistemy konechnoraznostnykh uraveniy, sootvetstvuyushchey zadache Dirikhle, s pomoshch'yu normal'noy posledovatel'nosti znakov [The Solution of Systems of Finite-Difference Equations Corresponding to the Dirichlet Problem, With Normal Number Sequence], DAN SSSR [Proc. Acad. Sci. USSR], 1958, 123, No. 3, pp. 407-409.

15. B.G. Gnedenko, Kurs teorii veroyatnostey [Course in the Theory of Probability], Moscow, 1954.

16. W. Sierpinski, Demonstration elementaire du theoreme de M. Borel sur les nombres absolument normaux et determination effective d'un tel nombre [Elementary Demonstration of Borel's Theorem on Absolutely Normal Numbers and Effective Determination of Such a Number], Bull. de la Societe mathematique de France [Bull. Math. Soc. France], 45 (1917), 127-132.

17. H. Lebesgue. Sur certaines demonstrations d'existence [Certain Existence Proofs], Bull. de la Societe Mathematique de France, 45, 1917, 132-144.

18. D.G. Champernowne, The Construction of the Decimals Normal in the Scale of Ten, J. London Math. Soc., 8, 1933, 254-260.

19. A.G. Postnikov, Issledovaniya po metodu trigonometricheskikh summ I.M. Vinogradova, Doktorskaya dissertatsiya, Mathematicheskiy

institut im. Steklova AN SSSR [Investigations by the Method of Trigonometrical Sums of I.M. Vinogradov, Doctorate Dissertation, Mathematical Institute Named for Steklov, Acad. Sci. USSR], 1956.

20. I.I. Shapiro-Pyatetskiy, O zakonakh raspredeleniya drobnykh doley pokazatel'noy funktsii [The Distribution Laws of Fractions of the Exponential Function], Izv. AN SSSR, ser. matem. [Bulletin of the Acad. Sci. USSR, Math. series], 15 (1951), 47-52.

21. A.G. Postnikov and I.I. Pyatetskiy, Normal'naya po Markovu posledovatel'nost' znakov i normal'naya tsepnaya drob' [Markov Normal Number Sequence and Normal Continued Fraction], Izv. AN SSSR, seriya matem. [Bulletin of Acad. Sci. USSR, Math. Series], 21 (1957), 729-746.

22. M.N. Martin, A Problem in Arrangements, Bull. of the Americ. Math. Soc., 40, No. 12, 1931.

23. N.M. Korobov, O normal'nykh periodicheskikh sistemakh [Normal Periodic Systems], Izv. Akad. nauk SSSR, ser. matem. [Bulletin of Acad. Sci. USSR, Math. Series, 16 (1952), 211-216.

24. Yu.N. Shakhov, Imitatsiya prosteyshchikh markovskikh protsessov [Imitation of the Simplest Markov Processes], Izv. AN SSSR, ser. matem. [Bulletin of the Acad. Sci. USSR, Math. Series], 23 (1959), 815-822.

25. N.M. Korobov, O vpolne ravnomernom raspredelenii i sovmestno normal'nykh chislakh [Completely Uniform Distribution and Consistently Normal Numbers], Izv. AN SSSR, ser. matem. [Bulletin of the Acad. Sci. USSR, Math. Series], 20 (1956), 649-660.

26. K. Mahler, On the Approximation of Logarithms of Algebraic Numbers, Phil. Trans. Roy. Soc., London A, 245, No. 898, 371 (1953).

27. L.P. Starchenko, Postroyeniye vpolne ravnomerno raspredelen posledovatel'nosti [Construction of Completely Uniformly D

tributed Sequence], DAN SSSR, 129, No. 3 (1959), 519-521.

28. I.P. Kubilyus and Yu.V. Linnik, Arifmeticheskoye modelirovaniye brounovskogo dvizheniya [Arithmetical Modeling of Brownian Motion], Izv. vysshikh zavedeniy, Matematika [Bulletin of Higher Institutions of Learning, Mathematics, 1959, No. 6 (13), 88-95.

29. I.I. Pyatetskiy-Shapiro, O raspredelenii drobnykh doley pokazatel'noy funktsii [Distribution of Exponential-Function Fractions], Uchenyye zapiski Moskovskogo gosudarstvennogo pedagogicheskogo instituta im. V.I. Lenina [Scientific Reports of the Moscow State Pedagogical Institute named for V.I. Lenin], CVIII, No. 2 (1957), 317-322.

30. S.S. Pillai, On Normal Numbers, Proc. Indian Acad. Sci., sec. A, 10 (1939), 13-15.

31. S.S. Pillai, On Normal Numbers, Proc. Indian Acad. Sci. sec. A, 12 (1940), 179-184.

32. I.Niven and H.S. Zuckerman, On the Definition of Normal Numbers, Pacific Journal of Mathematics, 1, No. 1, March, 1951.

33. I.W.S. Cassels, On a Paper of Niven and Zuckerman, Pacific Journal of Mathematics, 2 (1952), 555-557.

34. J.E. Maxfield, A Short Proof of Pillai's Theorem on Normal Numbers, Pacific Journal of Mathematics, 1 (1952), 23-24.

35. V.I. Romanovskiy, Diskretnyye tsepi Markova [Discrete Markov Chains], Moscow, 1949.

36. M.N. Korobov, Chisla s ogranychennym otnosheniyem i ikh prilozheniya k voprosam diofantovykh priblizheniy [Numbers With Limited Ratio and their Application to Problems of Diophantine Approximations], Izv. AN SSSR, ser. matem. [Bulletin of Acad. Sci. USSR, Math. Series], 19, (1956), 361-380.

37. L.P. Starchenko, O postroyenii posledovatel'nostey sovmestno

normal'nykh s dannoy [The Construction of Sequences Consistently Normal With a Given Sequence], Izv. AN SSSR, ser. matem. [Bulletin of Acad. Sci. USSR, Math. Series, 22 (1958), ?.7-370.

38. L.P. Starchenko, Ispravleniye k rabote "O postroyenii posledovatel'nostey sovmestno normal'nykh s dannoy [Corrections to Work entitled The Construction of Sequences Simultaneously Normal With a Given Sequence], Izv. AN SSSR, ser. matem. [Bulletin of Acad. Sci. USSR, Math. Series], 23 (1959), 635-636.

39. P. Khalmosh, Teoriya mery [Measure Theory], Moscow, 19?3.

40. I.P. Natanson, Teoriya funktsiy veshchestvennoy peremennoy [Theory of Functions of the Real Variable], Moscow, 195?

41. G.D. Birkhoff, Proof of a Recurrence Theorem for Strongly Transitive Systems. Proof of the Ergodic Theorem], Proc. Nat. Acad. Sci. USA, t. 17, 1931.

42. A. Khintchine. Zu Birkhoff Loesung des Ergodenproblems [Birkhoff Solution of the Ergodic Problem], Math. Annalen Bd. 107 (1932), 485-488.

43. F. Riez. Sur la theorie ergodique [Ergodic Theory], Commentarii math. Helvetici [Swiss Mathematical Commentaries], 17, F.3 (1944-1945), 221-239.

44. A.Ya. Khinchin, Tsepnyye drobi [Continued Fractions], Moscow, 1949.

45. K. Knopp, Mengentheoretische Behandlung einiger Probleme der diophantischen Approximationen und der transfiniten Wahrscheinlichkeiten [Group-Theory Treatment of Certain Problems of the Diophantine Approximations and Transfinite Probabilities], Math. Annalen, 85 (1926), 409-426.

46. ?. Reichenbach, Le fondement logique du calcul des probabilités [Logical Foundation of the Probability Calculus], Annales d.

l'Institut Henri Poincare [Annals of the Henri Poincare Institute],
7 (1937), 267-348.

47. A.H. Copeland, Admissible Numbers in the Theory of Geometrical Probability, Amer. Journ. of Math., 3 (19[illegible]), 153-162.

48. Dzh. M. Braun, Metody Monte-Karlo. Sbornik "Sovremennaya matematika dlya inzhenerov [Monte Carlo Methods. Collection entitled Contemporary Mathematics for Engineers], Moscow, 1958.

49. K. de Ley, E.F. Mur, K. Shennon and N. Shapiro, Vychislimost' na veroyatnostnykh mashinakh [Calculability on Probability Computers], Sbornik "Avtomaty" [Collection entitled Automatic Computers], Moscow, 1956.

50. A.H. Copeland and P. Erdos, Note on Normal Numbers, Bull. Americ. Math. Soc., 52 (1946), 857-860.

51. H. Davenport and P. Erdos, Note on Normal Decimals, Canadian; J. Math., 4 (1952) 58-63.

52. A.G. Postnikov. Kriteriy ravnomernogo raspredeleniya pokazatel'noy funktsii v kompleksnoy oblasti [Criterion of Uniform Distribution of Exponential Function in Complex Region], Vestnik Leningradskogo universiteta [Bulletin of Leningrad University], No. 13, No. 3, 81-88.

53. A.M. Polousev, O ra[illegible]mernom raspredelenii [illegible]emy [illegible]un[illegible]iy, yavlyayushcheysya resheniyem sistemy lineynykh ko[illegible]hn raznostnykh uraveniy perv[illegible]o poryadka (v pechati) [The Uniform Distribution of a System of Functions that is a Solution of a System of Linear Finite-Difference Equations of the First Order (at press)].

54. A.G. [illegible]kov, O [illegible]ichestve popadaniy drobnykh doley [illegible]kazatel[illegible]y funktsii na dannyy interval (v pechati) [Number of Occurrences of Exponential Fractions in a Given Interval (in printing)].

55. N.M. Korobov, Drobnyye doli pokazatel'nykh funktsiy [Fractions of Exponential Functions], Trudy Matematicheskogo instituta im. V.A. Steklova AN SSSR [Trans. of the Mathematical Institute named for V.A. Steklov Acad Sci. USSR], 38 (1951), 87-96.

56. N.M. Korobov, Nekotoryye mnogomernyye zadachi teorii diofantovykh priblizheniy [Certain Multidimensional Problems of the Theory of Diophantine Approximations], DAN SSSR, 84, No. 1, (1953), 13-16.

57. N.M. Korobov, Mnogomernyye zadachi raspredeleniya drobnykh doley [Multidimensional Problems of the Distribution of Fractions], Izv. AN SSSR, ser. matem. [Bulletin of the Acad. Sci., Math. Series], 17 (1953), 389-400.

58. A.M. Polousev, Mnogomernyy sluchay neuluchayemykh otsenok trigonometricheskikh summ s pokozatel'nymi funktsiyami [Multidimensional Case of Best Estimates of Trigonometric Sums with Exponential Functions], DAN SSSR, 104, No. 2 (1955), 186-189.

59. N.M. Korobov, Normal'nyye periodicheskiye sistemy i ikh prilozheniya k otsenke summ drobnykh doley [Normal Periodic Systems and Their Application to the Evaluation of Sums of Fractions], Izv. AN SSSR, ser. matem [Bull. of Acad. Sci. USSR, Math. Series], 15 (1951), 17-46.

60. R. Iglisch. Zum Aufbau der Wahrscheinlichkeitsrechnung [On the Construction of Probability Theory], Math. Ann., Vol. 107 (1932), 471-484.

61. H. Reichenbach, Axiomatik der Wahrscheinlichkeitsrechnung [Axioms of Probability Theory], Math. Zeit. [Mathematical Journal], Vol. 34, 1932, 568-619.

DISTRIBUTION LIST

DEPARTMENT OF DEFENSE	Nr. Copies
HEADQUARTERS USAF	
AFCIN-3D2	1
ARL (ARB)	1
OTHER AGENCIES	
CIA	1
NSA	6
DIA	9
AID	2
OTS	2
AEC	2
PWS	1
NASA	1
ARMY (FSTC)	3
NAVY	3
NAFEC	1
AFCRL (CRXLR)	1
PGE	12
SPECTRUM	1
RAND	1

MAJOR AIR COMMANDS	Nr. Copies
AFSC	
SCFDD	1
DDC	25
TDBTL	5
TDBDP	3
SSD (SSF)	2
APGC (PGF)	1
ESD (ESY)	1
RADC (RAY)	1
AFWL (WLF)	1
AFMTC (MTW)	1
ASD (ASYIM)	2